

Proceedings of the  
**Salford Postgraduate Annual  
Research Conference**  
(SPARC)  
2009

University of Salford





**University of Salford**  
A Greater Manchester University

**Proceedings of the Salford Postgraduate Annual Research Conference (SPARC) 2009**



This work is licensed under a Creative Commons Attribution-Non-Commercial-No Derivative Works 2.0 UK: England & Wales

You are free to copy and distribute this work or any part of it under the conditions that you give the original author(s) credit, that the work is not used for any commercial purposes, and that you do not change the work in any way.

For full details of this license, go to <http://creativecommons.org/licenses/by-nc-nd/2.0/uk/>

ISBN: 978-1-905732-88-3

Published by

The University of Salford  
Salford  
Greater Manchester  
M5 4WT  
United Kingdom

## Preface

Postgraduate research is the dynamo that drives innovative new work at good universities, and we are particularly proud of the work of our postgraduate students. SPARC brings together the best of this research across key fields that well express our university's strong themes of enquiry. The annual SPARC conference provides a dynamic forum for research students and supervisors to present and discuss new knowledge, fostering connections across disciplines that will open up areas for the future.

### **Professor Martin Hall**

Vice-Chancellor, University of Salford

Welcome to the 2009 proceedings of the Salford Postgraduate Annual Research Conference (SPARC). This selection of papers provides an excellent indication of the scope and scale of the two day interdisciplinary conference, and testifies to the quality and intellectual rigour of the research projects showcased there.

One of the defining characteristics of SPARC is its openness to all disciplines and research topics. This enables the participants to set the priorities of the conference, and allows for natural themes to emerge. This process has been carried through into this publication, which represents just under half of the papers given at the conference. This broad-ranging selection of thirty nine papers spans a wealth of disciplines – from health, medical sciences, management, finance, computing, education, and the built environment through to sociology, languages, literature, media, art history and cultural studies. But across the disciplinary boundaries of the papers certain themes are clearly identifiable – social inequality, community empowerment, changing organisational cultures, new kinds of creative economies, and the impact of digital technologies on learning, finance, business and cultural consumption to name but a few.

As well as facilitating conversations between disciplines, SPARC offers space for informal and academic networking between researchers from different universities, with presenters drawn from right across the North West region of the UK. Again, this publication offers a snapshot of that range: alongside Salford's researchers, there are contributions from postgraduates at the Universities of Liverpool, Cumbria, Chester, Central Lancashire, and from Liverpool John Moores and Manchester Metropolitan Universities. Building on this regional success, the conference committee have been encouraged to open up the 2010 and future SPARC events nationally and internationally. Providing an environment for cross - institutional engagement is just one of the ways in which the annual conference supports researcher development. It is an opportunity to develop individual research projects, giving presenters the chance to test out ideas on new audiences, receive valuable feedback from the academic community, make their work known and consequently raise their research profile. At the same time, the conference is also about professional development – allowing early career academics to develop all-important conference presentation skills, to practise the art of translating complex research ideas into a clear and concise format, and to communicate those ideas to audiences outside of their immediate specialism. This has valuable implications in terms of employability, encouraging the development of the kind of transferable skills that researchers are increasingly expected to demonstrate.

In developing the conference presentations into full papers for this publication, the contributors have expanded upon their research questions and findings in more detail, allowing readers privileged insight into a whole host of exciting and, in many cases, very new projects. The aim is for the proceedings to extend conversations well beyond the conference itself, by opening up new audiences and readerships for this emerging research. We are pleased to make the proceedings freely available online, and believe that the open access ethos of the unrestricted sharing of knowledge is very much in keeping with the spirit of the conference.

If you would like to find out more information about the annual conference and details other support for early career researchers offered by Salford's Graduate Studies team, please visit our website at <http://www.pg.salford.ac.uk>

**Dr Victoria Sheppard**

Research Skills Coordinator, University of Salford

## Contents

Issues in the Selection of Support Services for Outsourcing <i>Norbaya Ab Rahim, David Baldry, and Dilanthi Amaratunga</i> .....	<b>8</b>
1.5 T versus 3.0 T and 3D MPRAGE versus 3D MDEFT compared using Voxel- Based Morphometry <i>Ihssan Abdul-Kareem</i> .....	<b>16</b>
Examining and Improving the Limitations of the Gazis-Herman-Rothery Car- following Model <i>Hamid Al-Jameel</i> .....	<b>32</b>
Calibration of visual angle car following model based on real site traffic data <i>Jalal Al-Obaedi</i> .....	<b>45</b>
Changing Understanding of Emotion in an Educational Context <i>David Menendez Alvarez-Hevia</i> .....	<b>57</b>
Comparison of ADMS 4 and LiDAR in the Prediction of Atmospheric Boundary Layer <i>Vera Bachtiar, Chris Collier, Fay Davies</i> .....	<b>64</b>
Improving Integration in the Construction Industry through Knowledge Management in a Relational Contracting Setting: A Critical Success Factors Approach <i>Anis Sazira Bakri, Bingunath Ingirige, &amp; Dilanthi Amaratunga</i> .....	<b>74</b>
The Effect of Organisational Structure on Communication Behaviour in Libyan Public Organisations <i>Salh Bezweek and Charles Egbu</i> .....	<b>87</b>
E-Learning Adoption in Organisations: A Review of Existing Models and Practices <i>Ugochukwu C. Digwo, Dumebi O. Oderinde &amp; Mildred Brown-Houston</i> .....	<b>93</b>
E-business Adoption as a Form of Management Strategy: An e-Retailing Perspective <i>Ugochukwu C. Digwo &amp; Mildred Brown-Houston</i> .....	<b>107</b>
An investigation into the prevalence of <i>Toxoplasma gondii</i> infection in <i>Apodemus sylvaticus</i> in North Yorkshire <i>Nicole Dodd, Denise Thomasson, Elizabeth Wright, Jacqueline Hughes, &amp; Geoff Hide</i> .....	<b>124</b>
Lexical, Cultural, and Grammatical Translation Problems Encountered by Senior Palestinian EFL Learners at the Islamic University of Gaza, Palestine <i>Mohammed El Haj Ahmed</i> .....	<b>129</b>

Exploration Study of the Fishing Industry of Libya <i>Nagat Elmsallati, Prof. David Eaton</i> .....	<b>147</b>
The Impact of Contingent Factors on the Use of Performance Measurement Systems in the Banking Industry: The Case of Libya <i>Gumma Fakhri, Roger Pegum, Karim Menacere</i> .....	<b>153</b>
Appraisal for NHS doctors: a political wand or window dressing? (working paper) <i>Cao Fang, Frank Worthington &amp; Elaine Eades</i> .....	<b>170</b>
Current Usability Metrics for Mobile Computing Evaluation <i>Azham Hussain and Maria Kutar</i> .....	<b>185</b>
Mobility, Space, and Deserts: Urban Areas in Need of Precipitation <i>Bob Jeffery</i> .....	<b>191</b>
Factors hindering the adoption of E-learning in developing countries: Libya a case study - Methodological Perspectives <i>Abdulbasit S. Khashkhush, Professor David Eaton</i> .....	<b>204</b>
Time for Magic: An Investigation into Jamie Reid's Druid Ancestry <i>Vicki Maguire</i> .....	<b>211</b>
Defining Sustainability in Terms of Independent Filmmaking <i>Nico Meissner</i> .....	<b>221</b>
Social Enterprise Applications in an Urban Facilities Management Setting: Research Methodological Perspectives <i>Siti Uzairiah Mohd Tobi, Dilanthi Amaratunga</i> .....	<b>238</b>
Translators as Bridge-Builders between Cultures: Fact or Fiction? <i>Maria Nader</i> .....	<b>252</b>
Identifying and classifying stakeholders of post-disaster housing reconstruction projects in Sri Lanka <i>Nuwani Siriwardena, Richard Haigh &amp; Bingunath Ingirige</i> .....	<b>261</b>
Managing Disaster in Indonesia <i>Taufika Ophiyandri, Dilanthi Amaratunga &amp; Chaminda Pathirage</i> .....	<b>278</b>
The Retrospective Iconicity of 'Guerrillero Heroico' <i>Mike Parker</i> .....	<b>292</b>
Implications of the initial overt decisions on lung nodule detection from chest radiographs <i>Mariusz W. Pietrzyk</i> .....	<b>303</b>

<i>A Clockwork Orange: A Unique Challenge for the Translator</i> <i>Ines Radionovas</i> .....	<b>316</b>
Importance of quantifying orthotic correction of Adolescent Idiopathic Scoliosis: a retrospective study <i>Abdulla Rafiu, Stephen Hutchins</i> .....	<b>325</b>
A Framework for Geo-Spatial Data Sharing <i>Norzamni Salleh &amp; Professor Farzad Khosrowshahi</i> .....	<b>332</b>
Towards a Research Design to Investigate Assessment Practices in Built Environment Undergraduate Education <i>Lloyd Scott and Christopher Fortune</i> .....	<b>344</b>
The hospital as a physical space: accommodating children’s needs <i>Elham Sfandyarifard, Patricia Tzortzopoulos &amp; Monty Sutrisna</i> .....	<b>353</b>
Heaven, Hell and Complementarity: Opposition and Duality in late-Georgian and Regency England <i>Christine Simon</i> .....	<b>358</b>
An ethnographic study of the culture in a Diagnostic Imaging department (DID) – some initial findings <i>Ruth Strudwick</i> .....	<b>364</b>
The significance and co-dependency of ‘Robustness’ from the perspective of the Granting Authority in a PFI project environment <i>Gerald Sundaraj, Prof. David Eaton</i> .....	<b>374</b>
Remediated Swing and the Celestial Jukebox: new ‘friends’ for jazz? <i>Tom Sykes</i> .....	<b>390</b>
Using Virtual Learning Environment: Supporting Teaching and Learning, and Bridging the Digital Divide in Rural Communities in Ghana <i>Najim Ussiph</i> .....	<b>396</b>
An Uncertain Future for Urban Vegetation <i>Nicola Jane Wallbank &amp; Philip James</i> .....	<b>409</b>
A Voice for the People: the role of Social Media technologies in Community Empowerment and Regeneration <i>Eileen Wattam</i> .....	<b>417</b>
An Introduction to the Poetry of Selima Hill <i>Lucy Winrow</i> .....	<b>428</b>

# Issues in the Selection of Support Services for Outsourcing

**Norbaya Ab Rahim, David Baldry, and Dilanthi Amaratunga**, *School of the Built and Human Environment, University of Salford*

## Abstract

Outsourcing of support services (non-core activities) has now become an established trend for many UK organisations, both in the private and public sectors. The public healthcare sector is not excluded from this. Outsourcing, once used as a means to reduce costs, has now acquired strategic significance in organisations. Very little research has been carried out on the decision-making process of outsourcing in the public sector, particularly in healthcare. For many years, organisations have outsourced their support services to best-in-class service providers. This paper reviews the available literature on critical decision making in outsourcing support services, with particular emphasis on the selection process. It can be concluded that not only support services (non-core activities) are candidates for outsourcing; there are proposals to outsource core activities under certain circumstances. Understanding among staff about which activities in the organisation are core and non-core is also important. Therefore, it is crucial for organisations to make correct decisions when deciding which activities to outsource, as they have a substantial effect upon the future of the organisations.

**Keywords:** Decision making, outsourcing, selection, support services

## Introduction

Outsourcing is not a new concept (Russell and Taylor, 2003; Yang and Huang, 2000; Yang et al., 2007; Winkleman et al., 1993). For years businesses outsourced non-core services such as catering, accounting, information technology, and payroll, or used external consultants for these services. Even the type of decision that outsourcing represents is the same as the make-or-buy decision about products, processes, and facilities that organisations have been utilising for many years (Russell and Taylor, 2003). What has changed is the type and range of services being outsourced and the extent to which outsourcing has moved from a tactical to a strategic decision (Fill and Visser, 2000; Schniederjans and Zuckweiler, 2004). Outsourcing is increasingly becoming a strategic tool for many organisations, as this engages the unique talents of highly skilled suppliers in strategically important activities (Sanders et al., 2007). Momme and Hvolby (2002) argue that outsourcing should be considered as a dynamic process, and that any decisions concerning outsourcing should be in line with the strategic direction of the organisation.

Many studies on outsourcing decision-making have focused mainly on the private sector, especially manufacturing (Momme and Hvolby, 2002; McIvor, 2000; Fill and Visser, 2000; Schniederjans and Zuckweiler, 2004; Van de Water and van Peet, 2006; Barragan et al., 2003; Canez et al., 2000; Dekkers, 2000; Sisman and Satir, 2000; Venkatesan, 1992; Welch and Nayak, 1992). This paper proceeds from the premise that there is a lack of research on outsourcing decision making in the public sector, particularly in healthcare. Its analysis is part of the authors' doctoral research, which examines issues in the selection of support services for outsourcing in the public healthcare sector.

## Outsourcing

Chase et al. (2004: 372) define outsourcing as "an act of moving some of a firm's internal activities and decision responsibilities to outside providers". In the same vein, outsourcing is described as "the process by which a user employs the supplier, under a contract, to perform a function, which had previously been carried out in-house; and transfers to that supplier assets, including people and management responsibility" (Barrett and Baldry, 2003: 124). Furthermore, Lankford and Parsa (1999: 310) state that outsourcing is defined "as the procurement of products or services from sources that are external to the organisation". In the literature, the term *outsourcing* has been used interchangeably with the term *make-or-buy* (Van de Water and Van Peet, 2006; Canez et al., 2000; Venkatesan, 1992; Welch and Nayak, 1992; McIvor et al., 1997; McIvor and Humphreys, 2000; Fill and Visser, 2000; Probert,



1996). Make-or-buy is defined as “outsourcing the production process or parts of it” (Van de Water and Van Peet, 2006: 258). Outsourcing issues have been studied for many years by using nomenclature such as make-or-buy, vertical integration, or transaction cost analysis (Sanders et al., 2007). Outsourcing can be considered as a continuum. Mylott (1995) views outsourcing in terms of full outsourcing, selective outsourcing, and everything-in-between outsourcing.

### **Core and non-core activities**

Core activity is defined as “an activity central to the company successfully serving the needs of potential customers in each market. The activity is perceived by the customers as adding value and therefore being a major determinant of competitive advantage” (McIvor, 2000: 22). Meanwhile Hassanain (2005: 73) defines core activities as “those that are essential for achieving the objectives of the organisation”. On the other hand, non-core activities are those that are not critical to competitive advantage (Lonsdale, 1999: 176). In the same vein, Hassanain (2005: 73) adds that “non-core activities represent support services which are not part of core and are routinely performed.” As stressed by Quinn and Hilmer (1994), core competencies are those activities that offer long-term competitive advantage, whereas non-core activities are those that are not critical to the organisation’s competitive edge.

Although organisations possess their own core competencies, there exist inconsistencies and lack of clarity in the way in which the personnel within the organisation interpret these capacities (McIvor, 2003). McIvor further notes that the competencies that an organisation currently has may not be competitive in the future. This happens when the demands of customers change or when competitors are more capable in a particular activity. Similarly, activities that were not core in the past may become core in the future. For example, IT (Information Technology) was perceived as a non-core activity. However, with the emergence of computerised physician and electronic health records, IT is as core to patient care as any other hospital department (Ciotti and Pagnotta, 2005).

Hamel and Prahalad (1994) assert that it is dangerous for organisations to measure their competitiveness through price factor only, as this will lead to the erosion of their core competencies. Billi et al. (2004) point out that innovation, reputation, and strategic goals can all serve as a base in determining an organisation’s unique capabilities. It is not easy to differentiate between core and non-core activities (Jennings, 1997), and determining core competencies can be tricky and risky (Lankford and Parsa, 1999). A study carried by Lacity et al. (1995) suggests that not all activities considered strategic by an organisation are strategic in reality. Therefore, it is crucial for the organisation to carefully identify which activities are core and which activities are not.

In the healthcare sector, examples of core (also known as clinical) activities are pathology, radiology, physiotherapy, and pharmaceutical. Examples of non-core (non-clinical) activities are cleaning, security, cafeteria, laundry, car parking, landscape, and waste management.

### **Drivers for outsourcing**

There are numerous factors that encourage organisations to outsource. These include cost reduction (Fan, 2000; McIvor, 2003; Canez et al., 2000; Sanders et al., 2007; Fill and Visser, 2000; Gottschalk and Soli-Saether, 2005; Dole and Pinkard, 1993; Blumberg, 1998; Karyda et al. 2006); focus on core competencies (Fan, 2000; Hendry, 1995); improving quality (Fan, 2000; Canez et al., 2000); lack of internal skills, expertise, or capacity (Fan, 2000; Canez et al., 2000; Sanders et al., 2007; Fill and Visser, 2000); reduced time to market (Canez et al., 2000; McIvor, 2003); bandwagon effect of competitors’ outsourcing (Marshall et al., 2005); and technological change and advancement (Marshall et al., 2005; Sanders et al., 2007). Edgell et al. (2008) asserts that the drivers for outsourcing have been cyclical and tied to the economic cycle. During slower economic times, cost is the critical issue, whilst in good times other factors are of greater significance. However, it is important to note that the cost of outsourcing is not just the cost of the services, but also includes the costs of setting up relationships and monitoring and coordinating a service provider’s activities (Williamson, 1979).

Whilst cost and strategy may drive private organisations, the desire for the general well being of citizens may drive outsourcing by public organisations (Kremic et al., 2006). The instigation

to outsource can stem from policies created by government (Young, 2005); the requirements of managed care organisation or mergers (Moschuris and Kondylis, 2006); and agendas of elected officials (Young, 2005; Avery, 2000).

Although it can be argued that cost may not be the main driver for outsourcing in the public sector, there are studies that reveal that outsourcing in the public healthcare sector is motivated mainly by a desire to decrease costs, increase flexibility, share risks, focus on core competency, and increase the quality of services rendered (Yigit et al., 2007; Young, 2005). Yigit et al. (2007) further argue that outsourcing, when applied judiciously through cost and risk analysis, is a cost-effective approach that can be used by most hospitals. It should be noted that savings do not only occur through wage reductions, but also through changing work processes, rosters, and technology (Young, 2005). Although there are differences in the drivers of outsourcing between the private and public sectors, the desired benefits are nonetheless often similar (Kremic et al., 2006).

### **Advantages and disadvantages of outsourcing**

Yang and Huang (2000) emphasise that the most important considerations for any organisations contemplating outsourcing are the potential benefits to be gained. Amongst its key advantages, outsourcing allows organisations to concentrate on core activities (Fill and Visser, 2000; Hassanain and Al-Saadi, 2005; Yang et al., 2007; Welch and Nayak, 1992; Harland et al., 2005; Hendry, 1995; Moschuris and Kondylis, 2006); improve quality of service and performance (Hassanain and Al-Saadi, 2005; Fan, 2000; Moschuris and Kondylis, 2006); have flexibility in management (Yang et al., 2007; Welch and Nayak, 1992; Fan, 2000); reduce costs (Hassanain and Al-Saadi, 2005; Welch and Nayak, 1992; Harland et al., 2005; Fan, 2000; Moschuris and Kondylis, 2006); and provide creativity and innovation (Hassanain and Al-Saadi, 2005; Welch and Nayak, 1992). Outsourcing may therefore be an attractive method of improving an organisation's financial performance, especially in the short run (Gilley and Rasheed, 2000). However, precautions have to be taken, as long-term outsourcing contracts that have a feature of short-term savings can prove to be very expensive in the later stages (Okoroh et al., 2001; Arminas, 2009). Outsourcing in the healthcare sector can be a cost-effective way to provide health services because it brings in additional knowledge, expertise, and infrastructure. It can also render healthcare organisations more competitive and allow them to focus on their main goal of health service provision for patients (Mackey et al., 2004; Neil, 2004; Colona and McFaul, 2004). In sum, outsourcing can become a strategic tool used by healthcare management to control costs without affecting patient care (Colona and McFaul 2004).

On the other hand, outsourcing also has drawbacks. It can lead to low employee morale (Momme and Hvolby, 2002; Yang et al., 2007; Jennings, 2002; Young, 2005); loss of management control (Hassanain and Al-Saadi, 2005; Yang et al., 2007; Quinn and Hilmer, 1994); loss of critical skills and knowledge (Quinn and Hilmer, 1994; Jennings, 2002; Young, 2005); leakage of information security and confidentiality (Momme and Hvolby, 2002; Yang et al., 2007; Jennings, 2002; Young, 2005); imitation by service providers (Momme and Hvolby, 2002); weakening of innovative capability (Hassanain and Al-Saadi, 2005); and increased monitoring and management costs (Jennings, 2002; Young, 2005). In the same vein, it was reported by the Public Accounts Committee that mismanagement of government contracts with service providers could be costing taxpayers £290 million annually. In 2007–2008 alone, £12 billion was spent on service contracts, which cost approximately £240 million to manage (McNestrie, 2009). One of the most serious threats resulting from reliance on outsourcing is declining innovation by the outsourcer (Gilley and Rasheed, 2000). In addition, as suppliers gain knowledge of the product being manufactured, they may use that knowledge to begin marketing the product on their own (Pralhad and Hamel, 1990).

### **Selection of activities for outsourcing**

Deciding which activities to outsource is not an easy task. Fill and Visser (2000) are of the opinion that an organisation will tailor its decisions about what to outsource around its distinguishing features, with particular emphasis on the value and quality of its services. Critical in this context are the core business activities and competencies of the organisation. Because they provide competitive advantage, these should be kept in-house (Sislan and Satir, 2000; Quinn and Hilmer, 1994; Welch and Nayak, 1992; Lankford and Parsa, 1999; McIvor, 2009; Billi et al., 2004). Indeed, there are certain activities that should not be outsourced,

especially those that will affect the organisation's strategic core (Lonsdale, 1999) or those that share synergies with core activities (Tadelis, 2007). However, Beaumont and Sohal (2004) postulate that if the unique activities do not possess strategic advantage, an outsourcing partnership could be considered by the organisation. On the other hand, if the organisation manages to obtain specific access and skills related to those activities, they can be kept in-house.

A survey carried out by Fan (2000) reveals that most outsourced activities are peripheral in nature. Activities that have minimal or no impact on the competitive strategy (Sislan and Satir, 2000; Lankford and Parsa, 1999) can be outsourced. Activities with non-distinct capabilities are suitable candidates for outsourcing, as are matured, commodity-like, peripheral, tedious, and monotonous activities (Welch and Nayak, 1992; Venkatesan, 1992; Karyda et al., 2006; Bertolini et al., 2004; Beaumont and Sohal, 2004). Blumberg (1998) suggests outsourcing those activities that are not on par with those of the service providers in terms of productivity and efficiency. Moreover, McIvor (2009) claims that par-performance activities can also be considered for outsourcing when service providers can obtain a similar level of performance and no competitive advantage is created by the internal provider.

In the healthcare sector, outsourcing support (non-clinical) services is more common than outsourcing the core (clinical) functions (Sunseri, 1999). The decision regarding which activities to outsource in the healthcare sector can be difficult and politically charged (Billi et al., 2004).

On the basis of the core competency theory, Gordon and Zimmerman (2007) assert that outsourcing of core competency will not be acceptable. However, there are literatures that support the outsourcing of core activities. Van de Water and Van Peet (2006) argue that core activities will qualify for some form of outsourcing if the organisation lacks important strategic knowledge or skills. Hence, they suggest that organisations develop strategic alliances with the service providers. If the benchmarking exercise indicates that the organisation is not performing, the core activity should be strategically outsourced (McIvor, 1997).

Lacity and Hirschheim (1993) caution organisations to be extra vigilant when deciding which activities to outsource. This is because the perceived non-core or commodity activities may be very important to the running of the business; thus, extra care must be undertaken when deciding to outsource. On the other hand, the perceived strategic activities may become non-core or commodity services in the near future.

Any activities that have the potential of being competitively advantageous in the future must be considered seriously. Furthermore, the organisation should build this capability, even if outsourcing seems to be a very inviting option at that point in time (Sanders et al, 2007). Activities that do not furnish the organisation with a sustainable advantage and support its core activities directly should not be outsourced.

Whilst the existing literature suggests that non-core (support) activities are best suited for outsourcing, some authors maintain that organisations should outsource core activities under certain circumstances, such as when the organisation lacks important strategic knowledge or skills for a particular activity.

### **How to select these activities?**

Venkatesan (1992) suggests linking product differentiation, component families analysis, and manufacturing capability as a route to deciding which activities to outsource. Welch and Nayak (1992) adopt Venkatesan's work and develop a generic framework for outsourcing decision making. They argue that in addition to the issue of costing, organisations must consider strategic and technological factors when making outsourcing decisions. Pandey and Bansal (2003) propose three criteria for selecting which activities to outsource: criticality (strategic importance), stability (volatility), and simplicity (capability).

Gottfredson et al. (2005) likewise highlight three steps that can help in making objective outsourcing decisions: namely, identifying the core of the organisation's core activities; determining which capabilities are in need of outsourcing and what will be required to implement them; and identifying any potential loss of quality due to the outsourcing of capabilities. Quinn and Hilmer (1994) recommend that organisations to answer the following questions before deciding which activities to outsource:

1. What is the potential for obtaining a competitive advantage in this activity, taking account of transaction costs?
2. What is the potential vulnerability that could arise from market failure if the activity is outsourced?
3. What can be done to alleviate the vulnerability by structuring arrangements with suppliers to provide appropriate controls and allowing for necessary flexibilities in demand?

## Conclusions

This paper has highlighted issues in the selection of support services for outsourcing. It has noted that differentiating core and non-core activities is not an easy task for an organisation. Organisations have to be very careful when deciding which activities to retain in-house and which to outsource. There must be a consistent understanding amongst an organisation's staff regarding which activities are considered core and non-core. Therefore, a thorough decision-making exercise needs to be carried out to assure that outsourcing decisions do not result in a loss of competence in the near future.

## References

- Arminas, D. (2009) 'Bringing it in-house made good returns,' *FM World* 6 (7), p. 29.
- Avery, G. (2000) 'Outsourcing public health laboratory services: a blueprint for determining whether to privatize and how,' *Administration Review* 60 (4), pp. 330–337.
- Barragan, S., Cappellion, C., Dempsey, N., and Rothenberg, S. (2003) 'A framework for sourcing product development services,' *Supply Chain Management: An International Journal* (8) 3, pp. 271–280.
- Barrett, P. and Baldry, D. (2003) *Facilities Management: Towards Best Practice, 2<sup>nd</sup> ed.* Oxford: Blackwell Publishing.
- Beaumont, N. and Sohal, A. (2004) 'Outsourcing in Australia,' *International Journal of Operations and Production Management* 24 (7), pp. 688–700.
- Bertolini, M., Bevilacqua, M., Braglia, M., and Frosolini, M. (2004) 'An analytical method for maintenance outsourcing service selection,' *International Journal of Quality and Reliability Management* 21 (7), pp. 772–788.
- Billi, J.E., Pai, C.-W., and Spahlinger, D.A. (2004) 'Strategic outsourcing of clinical services: a model for volume-stressed academic medical centres,' *Health Care Management Review* 29 (4), pp. 291–297.
- Blumberg, D.F. (1998) 'Strategic assessment of outsourcing and downsizing in the service market,' *Managing Service Quality* 8 (1), pp. 5–18.
- Canez, L.E., Platts, K.W., and Probert, D.R. (2000) 'Developing a framework for make-or-buy decisions,' *International Journal of Operations and Production Management* 20 (11), pp. 1313–1329.
- Chase, R.B., Jacobs, F.R., and Aquilano, N.J. (2004) *Operations Management for Competitive Advantage, 10<sup>th</sup> Edition*. Boston: Irwin/McGraw-Hill.
- Ciotti, V. and Pagnotta, B. (2005) 'The other side of outsourcing,' *Healthcare Financial Management* 59 (2), pp. 38–41.
- Colona, J. and McFaul, W. (2004) 'Hospitals reconsider outsourcing key support and clinical departments,' *Health Care Strategic Management* 22 (7), pp. 11–14.
- Dekkers, R. (2000) 'Decision model for outsourcing and core competencies in manufacturing,' *International Journal Production Research* 38 (17), pp. 4085–4096.
- Dole, D. and Pinkard, L. (1993) 'The outsourcing source book,' *Journal of Business Strategy* 14 (3), pp. 52–56.
- Edgell, J., Meister, G.E., and Stamp, N. (2008) 'Global sourcing trends in 2008,' *Strategic Outsourcing: An International Journal* 1 (2), pp. 173–180.

- Fan, Y. (2000) 'Strategic outsourcing: evidence from British Companies,' *Marketing Intelligence and Planning* 18 (4), pp. 213–219.
- Fill, C. and Visser, E. (2000) 'The outsourcing dilemma: a composite approach to the make or buy decision,' *Management Decision* 38 (1), pp. 43–50.
- Gilley, K.M. and Rasheed, A. (2000) 'Making more by doing less: an analysis of outsourcing and its effects on firm performance,' *Journal of Management* 26 (4), pp. 763–790.
- Gordon, C. and Zimmerman, A. (2007) 'High-tech outsourcing: a benefit-cost framework,' *The American Economist* 51 (1), pp. 97–105.
- Gottfredson, M., Puryear, R., and Phillips, S. (2005) 'Strategic sourcing: from periphery to the core,' *Harvard Business Review* 83 (2), pp. 132–139.
- Gottschalk, P. and Solli-Saether, H. (2005) 'Critical success factors from IT outsourcing theories: an empirical study,' *Industrial Management and Data Systems* 105 (6), pp. 685–702.
- Hamel, G. and Prahalad, C.K. (1994) *Competing for the Future*. Boston: Harvard Business Press.
- Harland, C., Knight, L., Lamming, R., and Walker, H. (2005) 'Outsourcing: assessing the risks and benefits for organisations, sectors and nations,' *International Journal of Operation and Production Management* 25 (9), pp. 831–850.
- Hassanain, M.A. and Al-Saadi, S. (2005) 'A framework model for outsourcing asset management services,' *Facilities* 23 (1/2), pp. 73–81.
- Hendry, J. (1995) 'Outsourcing culture, community and networks: the hidden cost of outsourcing,' *European Management Journal* 13 (12), pp. 193–200.
- Jennings D. (2002) 'Strategic sourcing: benefits, problems and a contextual model,' *Management Decision* 40 (1), pp. 26–34.
- Karyda, M., Mitrou, E., and Quirchmayr, G. (2006) 'A framework for outsourcing IS/IT security services,' *Information Management and Computer Security* 14 (5), pp. 402–415.
- Kremic, T., Tukel, O.I., and Rom, W.O. (2006) 'Outsourcing decision support: a survey of benefits, risks and decision factors,' *Supply Chain Management: An International Journal* 11 (6), pp. 467–482.
- Lacity, M.C. and Hirschheim, R. (1993) 'The information systems outsourcing bandwagon,' *Sloan Management Review* 35 (1), pp. 73–86.
- Lacity, M.C., Willcocks, L.P., and Feeny, D.F. (1995) 'IT outsourcing: maximize flexibility and control,' *Harvard Business Review* May–June, pp. 94–93.
- Lankford, W.M. and Parsa, F. (1999) 'Outsourcing: a premier,' *Management Decision* 37 pp. 310–316.
- Lonsdale, C. (1999) 'Effectively managing vertical supply relationships: a risk management model for outsourcing,' *Supply Chain Management: An International Journal* 4 (4), pp. 176–183.
- Mackey, T.A., McNeil, N.O., and Klingensmith, K. (2004) 'Outsourcing issues for nurse practitioner practices,' *Nursing Economics* 22 (1), pp. 21–32.
- Marshall, D., Lamming, R., Fynes, B., and de Burca, S. (2005) 'The development of an outsourcing process model,' *International Journal of Logistics: Research and Applications* 8 (4), pp. 347–359.
- McIvor, R.T. (2000) 'A practical framework for understanding the outsourcing process,' *Supply Chain Management: An International Journal* 5 (1), pp. 22–36.
- McIvor, R.T. (2003) 'Outsourcing: insights from the telecommunications industries,' *Supply Chain Management: An International Journal*, 8(4) p380-394.
- McIvor, R.T. (2009) 'How the transaction cost and resource-based theories of firm inform outsourcing evaluation,' *Journal of Operations Management* 27 (1), pp. 45–64.
- McIvor, R.T. and Humphreys, P.K. (2000) 'A case-based reasoning approach to make or buy decision,' *Integrated Manufacturing Systems* 11 (5), pp. 295–310.

- McIvor, R.T., Humphreys, P.K., and McAleer, W.E. (1997) 'A strategic model for the formulation of an effective make or buy decision,' *Management Decision* 35 (2), pp. 169–178.
- McNestrie, A. (2009) 'Government slated over contract management,' *FM World* 6 (9), p. 6.
- Momme, J. and Hvolby, H. (2002) 'An outsourcing framework: action research in the heavy industry sector,' *European Journal of Purchasing and Supply Management* 8, pp. 185–196.
- Moschuris, S.J. and Kondylis, M.N. (2006) 'Outsourcing in public hospitals: a Greek perspective,' *Journal of Health Organization and Management* 20 (1), pp. 4–14.
- Mylott, T.R., III (1995) *Computer Outsourcing: Managing the Transfer of Information Systems*. Englewood Cliffs, NJ: Prentice Hall.
- Neil, R. (2004) 'Outsourcing that can work for your organisation,' *Material Management in Health Care* 13 (6), pp. 21–23.
- Okoroh, M.I., Gombera, P.P., John, E., and Wagstaff, M. (2001) 'Adding value to the healthcare sector: a facilities management partnering arrangement case study,' *Facilities* 19 (3/4), pp. 157–163.
- Pandey, V. and Bansal, V. (2004) 'A decision-making framework for IT outsourcing using the analytic hierarchy process,' *International Conference on Systemics, Cybernetics and Informatics*, Hyderabad, India.
- Prahalad, C.K. and Hamel, G. (1990) 'The core competence of the corporation,' *Harvard Business Review* July–August, pp. 79–91.
- Probert, D.R. (1996) 'The practical development of a make or buy strategy: the issue of process positioning,' *Integrated Manufacturing Systems* 7 (2), pp. 44–51.
- Quinn, J.B. and Hilmer F.G. (1994) 'Strategic outsourcing,' *Sloan Management Review* 35 (4), pp. 43–55.
- Russell, R.S. and Taylor, B.W. (2003) *Operations Management, 4<sup>th</sup> ed.* Upper Saddle River, NJ: Prentice-Hall.
- Sanders, N.R., Locke, A., Moore, C.B., and Autry, C.W. (2007) 'A multidimensional framework for understanding outsourcing arrangements,' *The Journal of Supply Chain Management: A Global Review of Purchasing and Supply* 43 (4), pp. 3–15.
- Schniederjans, M.J. and Zuckweiler, K.M. (2004) 'A quantitative approach to the outsourcing-insourcing decision in an international context,' *Management Decision* 42 (8), pp. 974–986.
- Sislan, E. and Satir, A. (2000) 'Strategic sourcing: a framework and a case study,' *The Journal of Supply Chain Management: A Global Review of Purchasing and Supply* 36 (3), pp. 4–11.
- Sunseri, R. (1999) 'Outsourcing on the outs,' *Hospitals and Health Networks* 73 (10), pp. 46, 48, 50–52.
- Tadelis, S. (2007) 'The innovative organization: creating value through outsourcing,' *California Management Review* 50 (1), pp. 261–277.
- Van de Water, H. and van Peet, H.P. (2006) 'A decision support model based on the analytic hierarchy process for the make or buy decision in manufacturing,' *Journal of Purchasing and Supply Management* (12) 5, pp. 258–271.
- Venkatesan, R. (1992) 'Strategic outsourcing: to make or not to make,' *Harvard Business Review* 70 (6), pp. 98–107.
- Welch, J.A. and Nayak, P.R. (1992) 'Strategic sourcing: a progressive approach to make or buy decision,' *Academy of Management Executive* 6 (1), pp. 23–30.
- Williamson, O. (1979) 'Transaction cost economics: the governance of contractual relations,' *Journal of Law and Economics* 22 (2), pp. 233–261.
- Winkleman, M., Dole D., Pinkard, L., and Molloy, J. (1993) 'The outsourcing source book,' *Journal of Business Strategy* 14 (3), pp. 52–56.

- Yang, C. and Huang, J.B. (2000) 'A decision model for IS outsourcing,' *International Journal of Information Management* 20 (3), pp. 225–239.
- Yang, D-H., Kim S. Nam, C., and Min, J.-W. (2007) 'Developing a decision model for business process outsourcing,' *Computers and Operations Research* 34 (12), pp. 3769–3778.
- Yigit, V., Tengilimoglu, D., Kisa, A., and Younis, M.Z. (2007) 'Outsourcing and its implications for hospital organizations in Turkey,' *Journal of Health Care Finance* 33 (4), pp. 86–92.
- Young, S. (2005) 'Outsourcing in the health sector: the interplay of economics and politics,' *International Journal of Public Sector Management* 18 (1), pp. 25–36.

# 1.5 T versus 3.0 T and 3D MPRAGE versus 3D MDEFT compared using Voxel-Based Morphometry

**Ihssan Abdul-Kareem**, *School of Health Sciences, Division of Medical Imaging and Radiotherapy, University of Liverpool*

## Abstract

**Objective** using VBM to unveil whether scanning same subjects at different field strengths or with different pulse protocols may or may not introduce gray and white matter brain differences that may be falsely interpreted as biological ones.

**Subjects and Methods** 10 healthy adults (5 males and 5 females) with an average age of 31 years were scanned thrice, once at 1.5T using MPRAGE and twice at 3.0T using MPRAGE and MDEFT pulse sequences. Two VBM comparisons of gray and white matter were made using SPM5: 1.5T MPRAGE versus 3.0T MPRAGE and 3.0T MPRAGE versus 3.0T MDEFT.

**Results** several gray and white matter differences were detected over the whole brain in both comparisons. These differences were more prevalent in 1.5T versus 3.0T comparison as well as in 3.0T less than 1.5T and MDEFT less than MPRAGE contrasts.

**Conclusions** as we compared same subjects in both comparisons, these gray and white differences observed in our study are technical non-biological differences that are due to differences in signal to noise ratio, contrast to noise ratio and artefacts. We propose that combining scans obtained at different field strengths or with different pulse sequences in MR studies is not without the bias of introducing technical variabilities that might adversely affect validity and reliability of these studies. The same precaution should be considered when comparing results obtained from studies using different field strengths and/or different pulse sequences.

**Key words:** Scanner comparison, 1.5 Tesla, 3.0 Tesla, MPRAGE, MDEFT, VBM.

## 1.0 Introduction

With the increase use of MRI scanners operating at field strength higher than 1.5 Tesla- such as 3.0 Tesla- in research and clinical fields, there is an increasing demand to unveil whether the results obtained from a 3.0 T scanner are comparable to those of a 1.5 T scanner. From the 'MR physics' point of view, there are many differences between scans obtained at these field strengths. The shift to a higher field strength [1-4] offer the advantages of shorter examination times, improved chemical selectivity, increased signal to noise ratio (SNR) and higher image contrast which together with the improved SNR lead to higher contrast to noise ratio (CNR). However, it has the disadvantages of decreasing T1 tissue contrast as well as increasing acoustic noise, chemical shift artefacts and susceptibility-induced geometric distortion inversely affecting image quality. These latter effects have been used to advantage in MR spectroscopy and fMRI[5, 6], respectively. Finally, the use of high field strengths increases the likelihood of exceeding regulatory limits of specific absorption rate beside the uncertainty of MR compatibility and safety profiles when using these field strengths.

Several studies addressed the issue of comparing scans obtained at 1.5 T and 3.0 T field strengths but, specifically the following is going to consider estimated volumes. A structural study that used manual tracing reported that hippocampal volumes measured in normal subjects were not different between scans obtained at 1.5 T and 3.0 T field strengths[7]. Functional studies reported that the number of activated voxels were significantly higher at 3.0 T[8] and that the spectra in phantoms demonstrated improved resolution at 3.0 T compared to 1.5 T[9]. In clinical practice, 3.0 T scanners have been widely used since 1999[1] and, in many centres, replaced 1.5 T scanners creating new challenges for neuroradiologists. Many clinical studies reported that increasing field strength would increase diagnostic efficacy: It



was found that imaging at 3.0 T would permit more lesion detection in acute optic neuritis[10], multiple sclerosis[11] and produce higher contrast between tumor and normal brain following the administration of gadolinium as compared to 1.5 T[12, 13].

Combining structural images obtained from scanners of different field strengths may be justified as a time saving procedure allowing researchers to obtain large number of images in a short time. However, variations in imaging parameters through the use of different field strengths can result in changes in contrast properties of resulting images that are independent of the underlying tissue but rather reflect the physics of the imaging process[14]. It is, therefore, of extreme importance to know whether these changes in contrasts properties may introduce differences and variability in morphometric brain measures that do not reflect biological effects but rather technological ones and thus should be minimized. This is especially true for scanners operating at 1.5T and 3.0T field strengths as they are the most commonly used scanners in clinical MRI centres.

Apart from field strength, it is also important to rule out whether the use of different pulse protocols may introduce non-biological variations in scans. The dramatic development in the field of neuroimaging both in research and clinical fields led to an increase demand of obtaining three dimensional (3D) T1-weighted structural brain scans with high resolution, CNR, SNR and a short acquisition time. Such structural scans are widely used in research and clinical practices. The most commonly used pulse sequences for obtaining scans with such characteristics are MPRAGE (Magnetization-Prepared Rapid Gradient-Echo imaging)[15, 16] and the relatively new, MDEFT (Modified Driven Equilibrium Fourier Transform) [17, 18].

MPRAGE sequence has been optimized for structural brain imaging at field strengths of 1.5 T[19], 2.0T[20] and 3.0 T[21]. MDEFT sequence, on the other hand, has been widely used for T1 weighted structural brain imaging especially at high field strengths[17] as it is suggested to provide a good contrast between gray matter (GM) and white matter (WM). Basically, It has been used for anatomical brain imaging at 4.0T[18], 8.0T[22] and, more recently, optimized for 1.5 T and 3.0 T field strengths[23]. Both MPRAGE and MDEFT are gradient echo sequences that use 3D (FLASH) (Fast Low Angle Shot) imaging technique for data acquisition[24] with very short repetition time, echo time and very low flip angle. Both were used in fMRI studies as a reference for functional scans (which usually display low spatial resolution and reduced tissue contrast)[25, 26], in volumetric studies[27], and – particularly relevant to our study - both MPRAGE[28-30] and MDEFT[31-33] had been widely used in morphometric studies such as voxel-based morphometry (VBM) as the excellent GM/WM contrast and high spatial resolution they provide can be used to advantage allowing accurate voxel-wise comparison.

In clinical practice, the rapidity of acquisition and the capability of 3D rendering make MPRAGE an ideal sequence[34]. It has been shown to be effective in improving the sensitivity of lesion detection in patients with multiple sclerosis[10, 35], brain tumors[12, 13, 36], as a complementary sequence to improve inner ear and facial nerve imaging[37, 38], as the technique of choice for the examination of ENT tumors near the base of skull[39] and in simplifying and enhancing brain tissue segmentation[40].

To date, the previously published studies pointed out that MPRAGE is less favourable than MDEFT in high field strengths: at lower field strengths (<3.0 T), MPRAGE provide good GM/WM contrast; however, it performs more poorly at higher field strengths due to the longer T1 values and here the MDEFT was proposed to have advantage[17, 45]. Another advantage of MDEFT over MPRAGE is its relative insensitivity to inhomogeneities in  $B_1$  field[3]. A study [46]that compared T1-wighted 3D images obtained at 3.0 T using 3 sequences (FLASH, MPRAGE and MDEFT) in humans and phantoms reported that MDEFT has the highest CNR efficiency making this sequence ideal for applications that require WM-GM tissue segmentation such as the VBM. The authors suggested that at 3.0 T, it is preferable to implement MDEFT with parallel-imaging or partial k-space sampling methods to reduce the scan time than to use MPRAGE. These findings were replicated by another 3.0 T study[47] conducted on post-mortem brains reporting that MDEFT sequence is optimum for GM/WM contrast with contrast values superior to those of MPRAGE. Despite these advantages, the MDEFT sequence has not been used widely in research and clinical practices possibly due to its relative recent introduction and the lack of sufficient baseline data. As this sequence was originally designed for high field strengths, a comparison between it and the MPRAGE sequence in high field strength such as the 3.0 T would be valuable.

To unveil whether scanning same subjects at different field strengths or with different pulse protocols may or may not produce technical GM and WM differences that might be falsely

interpreted as biological ones, we use VBM technique[48] as it provides an objective examination of the whole brain. We perform two comparisons: first, we use VBM to compare scans of healthy subjects obtained at two field strengths (1.5 T and 3.0 T) using identical pulse protocol (T1-3D MPRAGE). In the second, we use same technique for comparing scans of same subjects but obtained using two different pulse protocols (T1-3D MPRAGE and T1-3D MDEFT) both at a field strength of 3.0 T.

## 2.0 Materials and methods

### 2.1 Subjects

10 right handed subjects of both sexes (5 males and 5 females) with a mean age of 31 years and a standard deviation of 6 years were participated in this study. All subjects were in good neurological and psychological health as assessed by a complete medical questionnaire including medical history. All subjects have given fully informed written consent to participate in this study that was approved by the local research ethics committee.

### 2.2 MRI data acquisition

High resolution 3D T1-weighted MR images were acquired on a SYMPHONY 1.5 T and TRIO 3.0 T whole body MRI imaging system (SIEMENS, MAGNETOM, GERMANY). Each subject was scanned three times, once on the 1.5 T scanner and twice on the 3.0 T scanner. In five subjects, all three scans were conducted at the same day and in the rest, scanning was conducted at two separate days with an average of one month duration between examinations during which subjects have not suffered from any neurological or psychological illness and have not taken any specific drug or underwent surgical operations.

On the 1.5 T scanner, subjects were examined using a 3D MPRAGE pulse sequence (TR = 2040 ms, TE = 3.93 ms, TI = 1100 ms, flip angle =  $15^\circ$ ) using 8 channels –receive only-head coil and a bandwidth of 130 Hz/Px. A field of view (FOV) of 25.6 cm was used with an acquisition matrix of 256 X 224 producing 176 contiguous T1-weighted sagittal slices with slice thickness of 1 mm. Acquired images were of voxel size 1.0 X 1.0 X 1.0 mm, the acquisition time was 7 minutes and 38 seconds and the number of excitations was 1.0.

On the 3.0 T scanner, subjects were examined using 3D MPRAGE and 3D MDEFT pulse sequences. All images were obtained using 8 channels –receive only-head coil and whole body transmit coil. For the 3D MPRAGE sequence, the parameters were (TR = 2300 ms, TE = 4.37 ms, TI = 1100 ms, flip angle =  $8^\circ$ ) with a bandwidth of 190 Hz/Px. The FOV was 20 cm with an acquisition matrix of 320 X 320 producing 192 contiguous sagittal slices with slice thickness of 1 mm. Acquired images were of voxel size 0.6 X 0.6 X 1.0 mm, the acquisition time was 12 min and 18 s and the number of excitations was 1.0. For the 3D MDEFT sequence, the parameters were (TR = 7.92 ms, TE = 2.48 ms, TI = 910 ms, flip angle =  $16^\circ$ ) with a bandwidth of 195 Hz/Px. A FOV of 25.6 cm with an acquisition matrix of 256 X 240 producing 176 contiguous sagittal slices with slice thickness of 1 mm. Acquired images were of voxel size 1.0 X 1.0 X 1.0 mm, the acquisition time was 12 minutes and 51 seconds and the number of excitations was 1.0.

In short, the 3D MPRAGE pulse sequence consists of a three-step cycle [15]: (a) magnetization preparation for contrast control (a  $180^\circ$  inversion pulse inverts the magnetization yielding the T1 contrast); (b) data acquisition using a FLASH sequence[24] with very short repetition time, echo time and very low flip angle; and (c) magnetization recovery for additional contrast control. An additional phase-encoding gradient in the slice selection direction is applied for 3D Fourier transform imaging.

For the 3D MDEFT sequence, we used the optimized method described by [23]. It also consist of 3 step cycle: (a) a global saturation pulse to null both the longitudinal and transverse magnetization; (b) after a recovery time delay, a  $180^\circ$  pulse is applied to invert the longitudinal magnetization; (c) after an inversion recovery delay, a 3D imaging technique (3D FLASH) (Fast Low Angle Shot) is used to acquire image data. The sequence is repeated a number of times in order to build up the whole imaging data matrix, with each repeat acquiring data for different regions of the 3D Kspace volume.

### 2.3 VBM analysis

Voxel-based morphometric analysis of data was performed using SPM5 software (Statistical Parametric Mapping, Wellcome Department of Imaging Neuroscience, University College London, UK; available at <http://www.fil.ion.ucl.ac.uk/spm>) implemented in MATLAB & SIMULINK, version 7.4.0.287 - R2007a., The MathWorks, Inc., Natick, MA, USA.

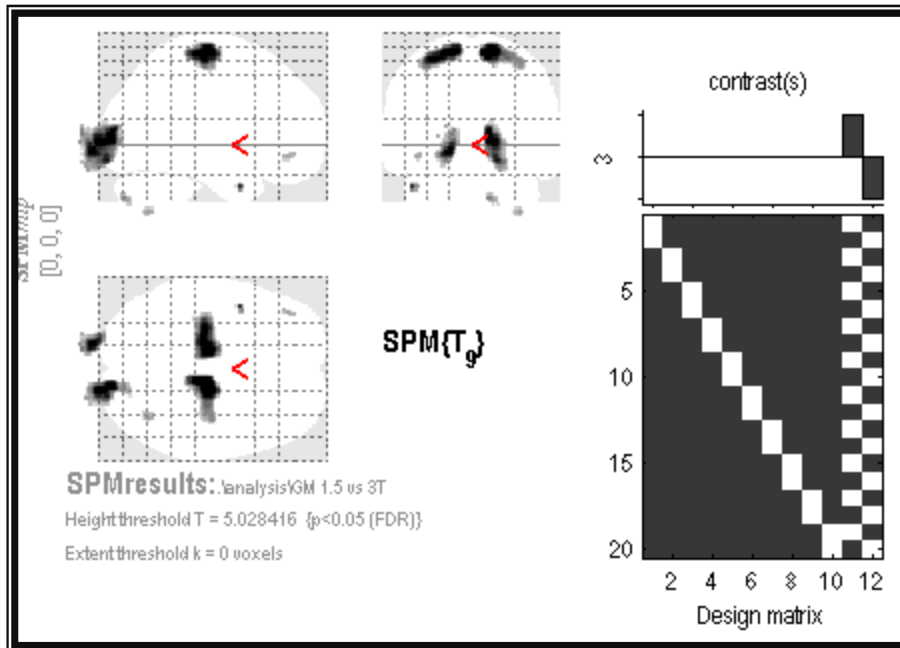
Generally speaking, VBM technique is based on four main steps: first, spatial normalization of structural images to the same stereotactic space which involves estimating the optimum 12-parameter affine transformation that maps individual MRI images to a template and correcting for global nonlinear shape differences, which are modeled by a linear combination of smooth spatial basis functions. This method does not attempt to match every cortical feature exactly, but merely corrects for global brain shape differences. Second, segmentation of the normalized images into tissue types (GM and WM). this is achieved by combining a priori probability maps or "Bayesian priors" with a mixture model cluster analysis which identifies voxel intensity distributions of particular tissue types. The segmentation step also incorporates an image intensity non-uniformity correction to account for smooth intensity variations caused by different positions of cranial structures within the MRI coil. Third, smoothing segmented images which is done by convolving with a 3D Gaussian kernel defined by its full-width-at-half-maximum (FWHM). This step has the advantages of (a) ensuring that each voxel in the images contains the average amount of GM or WM from around the voxel, (b) rendering the data more normally distributed thus increasing the validity of parametric statistical tests (c) potentially increasing SNR and (d) compensating for the inexact nature of the spatial normalization thus reduces the effective number of statistical comparisons. Fourth, performing a statistical analysis to localize significant differences between experimental groups that will be displayed in a statistical parametric map (SPM).

Unlike older versions of SPM such as SPM99 and SPM2, the spatial normalization step is not required in SPM5 in which the segment button allows this step to be achieved using a different objective function which does not rely on a simple relationship between the intensities of a pair of images. By using unified segmentation, tissue classification, bias correction and non-linear warping could be combined within the same framework[49]. Prior to segmentation, images were realigned to the anterior commissure to achieve better registration. Images were then segmented into two tissue classes: GM and WM and then smoothed using an isotropic Gaussian Kernel with FWHM of 10 mm. Finally, the smoothed images were statistically analyzed using an SPM5 group comparison. We performed a paired *t*-test to test for regional differences in GM and WM between two conditions (brain scans generated at 1.5 T versus those generated at 3.0 T, both obtained using MPRAGE sequence) and (brain scans obtained using MPRAGE versus those obtained using MDEFT pulse sequences, both generated at 3.0T field strength). We assumed that scans have equal variance as we compared scans of same subjects. Results were thresholded at  $P$  (FDR- corrected)  $< 0.05$ .

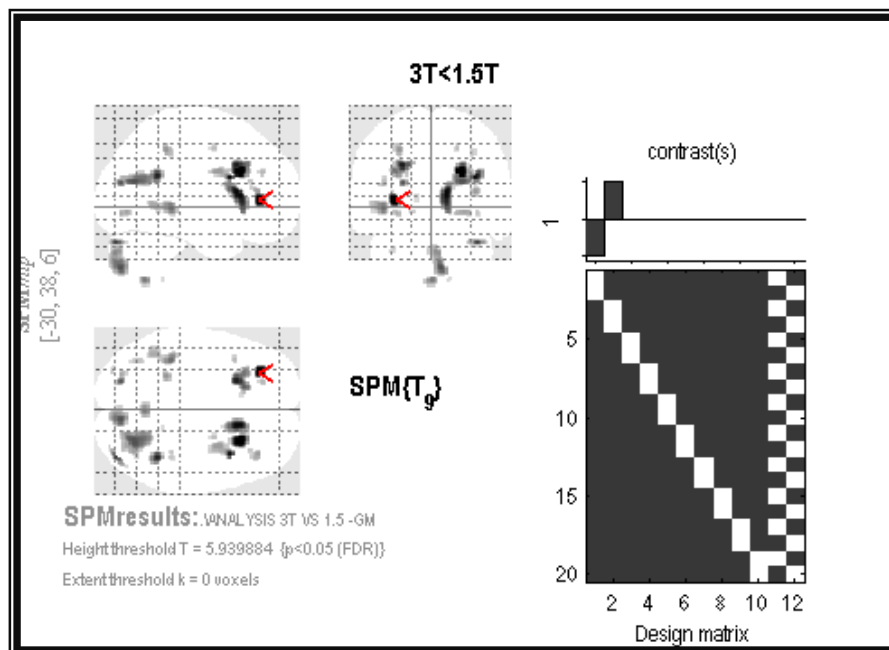
### 3.0 Results

For each comparison, we defined two contrasts: a) 3.0T greater than 1.5T and 3.0T less than 1.5T and b) MDEFT greater than MPRAGE and MDEFT less than MPRAGE. We applied these contrasts for GM and WM comparisons which ultimately resulted in having eight (SPM. mat) files being analyzed. We used the Talairach client available online at <http://www.talairach.org/client.html> to delineate the location of each significant cluster. For GM differences, localization of regions was according to the nearest gray matter voxel while for WM differences, localization was restricted to a single point.

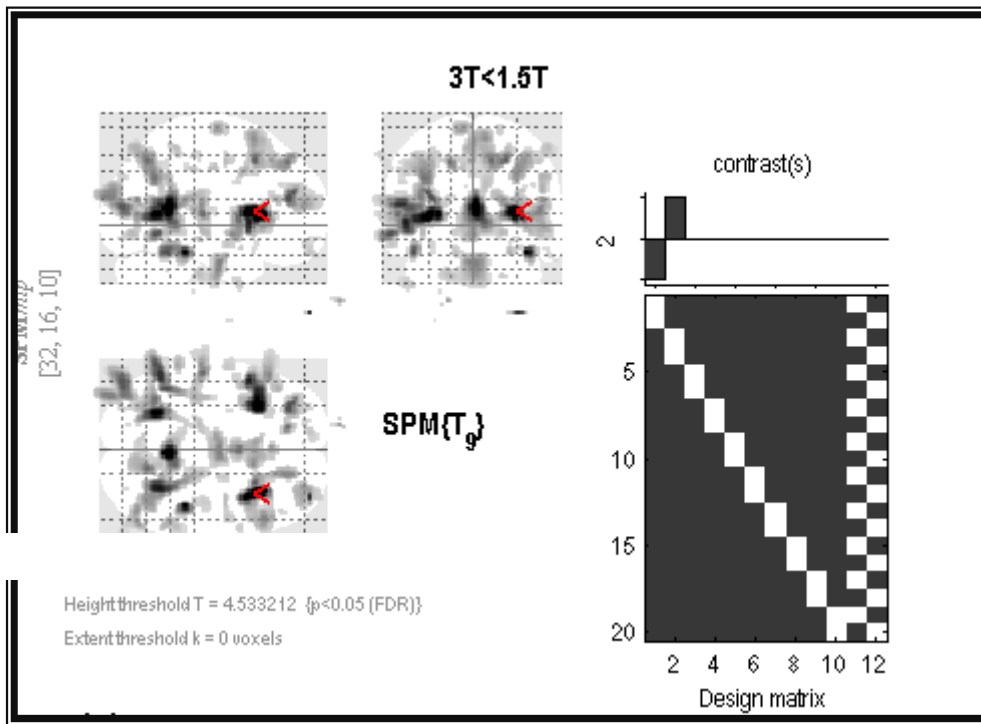
Results show several GM and WM differences in both comparisons. These differences were more prevalent in 1.5T less than 3.0T and MDEFT less than MPRAGE contrasts in both GM and WM. Furthermore, these differences were more prevalent when comparing 1.5T versus 3.0 scanners. Differences were observed in many brain regions including cerebral cortical areas (frontal, parietal, temporal limbic and occipital lobes), subcortical and sublobar areas (insula, caudate nucleus) and cerebellum. Figures (1-5) show the resulted statistical parametric maps of comparisons. Tables (1-8) show Talairach coordinates[50], Z scores and anatomical regions for significant clusters including Brodmann areas[51] for GM regions. These data reveal clearly the existence of several GM and WM differences in both (1.5T versus 3.0T) and (MDEFT versus MPRAGE) comparisons.



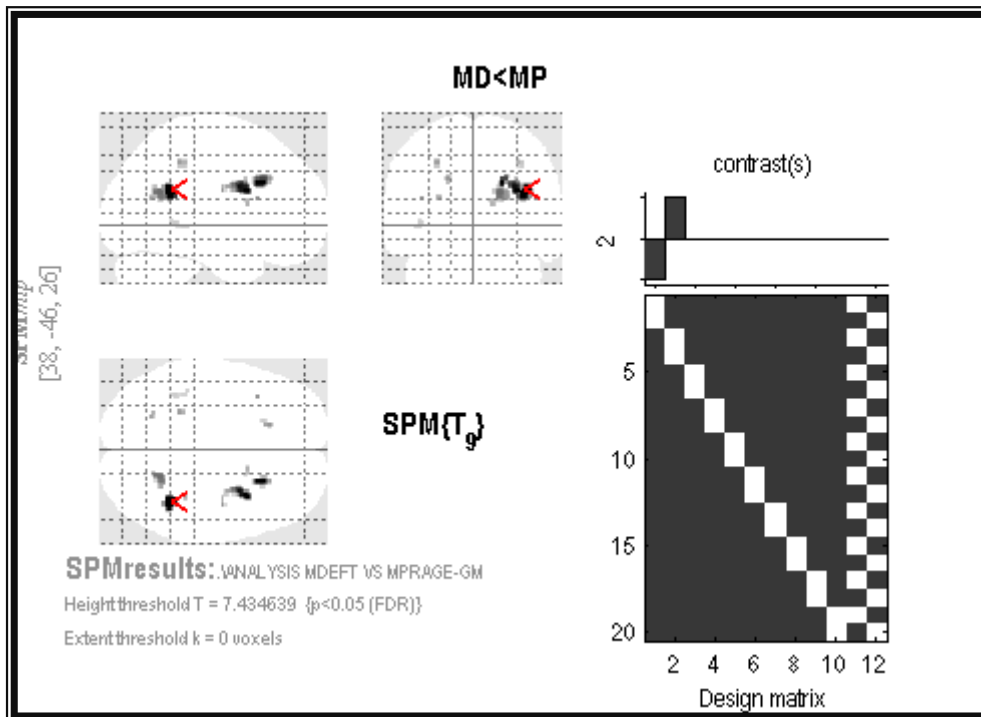
**Figure 1.** Statistical parametric maps, in sagittal, coronal and axial projections, showing clusters of increased gray matter in 3.0 Tesla MPRAGE scans relative to 1.5 Tesla MPRAGE scans.



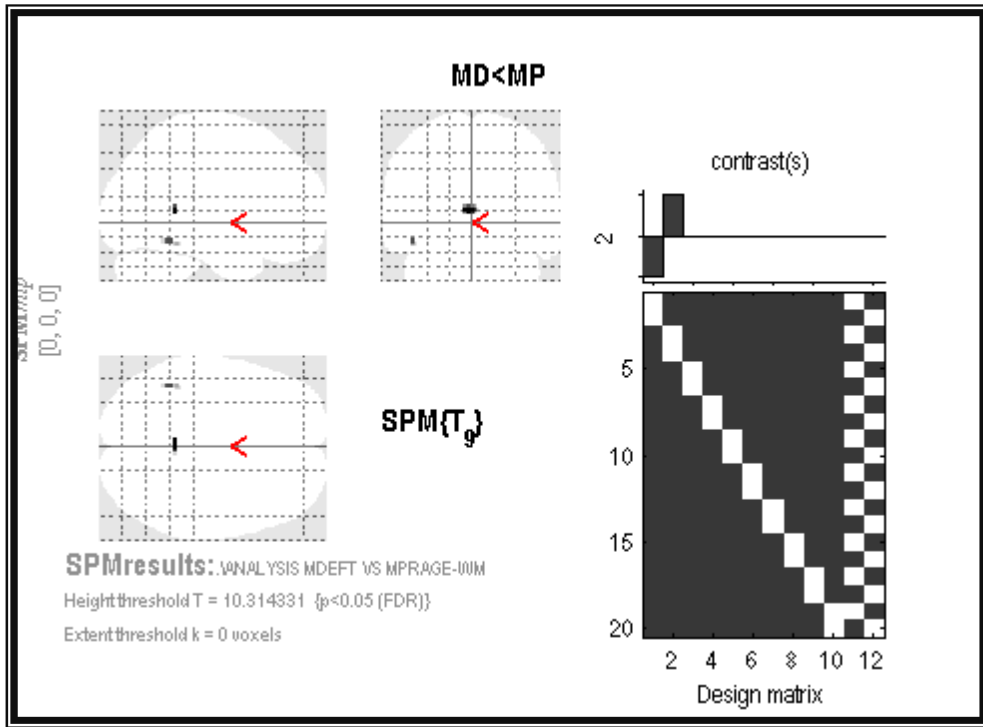
**Figure 2.** Statistical parametric maps, in sagittal, coronal and axial projections, showing clusters of decreased gray matter in 3.0 Tesla MPRAGE scans relative to 1.5 Tesla MPRAGE scans.



**Figure 3.** Statistical parametric maps, in sagittal, coronal and axial projections, showing clusters of decreased white matter in 3.0 Tesla MPRAGE scans relative to 1.5 Tesla MPRAGE scans.

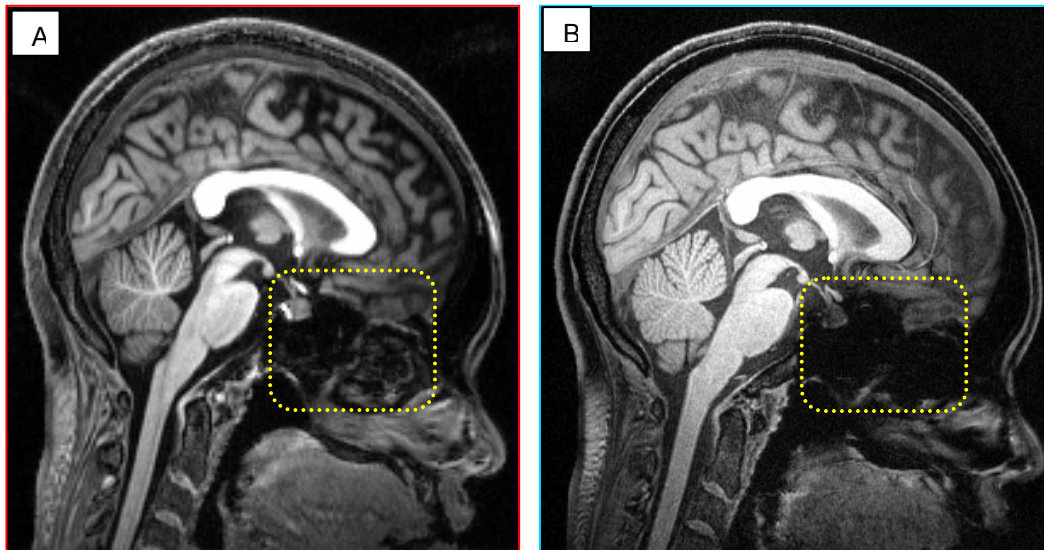


**Figure 4.** Statistical parametric maps, in sagittal, coronal and axial projections, showing clusters of decreased gray matter in 3.0 Tesla MDEFT scans relative to 3.0 Tesla MDEFT scans.



**Figure 5.**  
Statistical

parametric maps, in sagittal, coronal and axial projections, showing clusters of decreased white matter in 3.0 Tesla MDEFT scans relative to 3.0 Tesla MDEFT scans



**Figure 6.** Susceptibility artefacts in the regions of nasal cavity and paranasal air sinuses (dotted box) in midline sagittal images of the same subject. A. 3.0T MDEFT sequence. B. 3.0T MPRAGE sequence. The artefacts are more prominent with the MPRAGE sequence.

**Table 1.** 3.0T greater than 1.5T GM differences.

<b>X</b>	<b>Y</b>	<b>Z</b>	<b>*Z score</b>	<b>Region</b>	<b>BA</b>
10	-20	66	7	Right cerebrum, frontal lobe, precentral gyrus	6
-12	-16	60	6.2	Left cerebrum, frontal lobe, precentral gyrus	6
-44	6	30	6.3	Left cerebrum, frontal lobe, inferior frontal gyrus	9
-40	42	-8	6.1	Left cerebrum, frontal lobe, middle frontal gyrus	47
12	-9	4	5.2	Right cerebrum, sublobar, thalamus, ventral lateral nucleus	
-16	-104	-4	5	Left cerebrum, occipital lobe, cuneus	18
14	-60	-38	5	Right cerebellum, posterior lobe, cerebellar tonsil	

\*P<0.05 (FDR corrected)

BA, Brodmann area

**Table 2.** 3.0T greater than 1.5T WM differences.

<b>X</b>	<b>Y</b>	<b>Z</b>	<b>Z score</b>	<b>Region</b>
16	2	-30	2.5	Right cerebrum, limbic lobe, uncus
-22	-68	-38	3.5	Left cerebellum, posterior lobe, cerebellar tonsil
32	-44	-42	3.4	Right cerebellum, posterior lobe, cerebellar tonsil

\*P<0.05 (FDR corrected)

**Table 3.** 3.0T less than 1.5T GM differences.

<b>X</b>	<b>Y</b>	<b>Z</b>	<b>*Z score</b>	<b>Region</b>	<b>BA</b>
-30	38	6	5.9	Left cerebrum, frontal lobe, middle frontal gyrus	10
36	38	12	4.3	Right cerebrum, frontal lobe, middle frontal gyrus	10
-22	40	28	4.2	Left cerebrum, frontal lobe, superior frontal gyrus	9
30	4	28	5.8	Right cerebrum, frontal lobe, precentral gyrus	6
-22	20	34	4.3	Left cerebrum, frontal lobe, middle frontal gyrus	8
-28	26	28	3.7	Left cerebrum, frontal lobe, middle frontal gyrus	9
-40	18	18	3.6	Left cerebrum, frontal lobe, middle frontal gyrus	46
22	0	48	3.2	Right cerebrum, frontal lobe, middle frontal gyrus	6

34	-34	44	4.3	Right cerebrum, parietal lobe, inferior parietal lobule	40
-26	38	46	5.6	Left cerebrum, parietal lobe, post central gyrus	3
-40	-34	0	4.7	Left cerebrum, temporal lobe, subgyral	
-44	-50	2	4.2	Left cerebrum, temporal lobe, middle temporal gyrus	37
-38	-62	22	5.3	Left cerebrum, temporal lobe, middle temporal gyrus	39
38	-46	26	4.8	Right cerebrum, insula	13
-4	22	28	5.4	Left cerebrum, limbic lobe, anterior cingulate gyrus	24
12	22	10	4.2	Right cerebrum, limbic lobe, anterior cingulate gyrus	24
26	-64	22	5.8	Right cerebrum, limbic lobe, posterior cingulate gyrus	31
-14	24	-2	4.3	Left cerebrum, head of caudate nucleus	
-10	22	8	3.7	Left cerebrum, body of caudate nucleus	
38	-34	-4	3.6	Right cerebrum, tail of caudate nucleus	
28	-72	-2	3.2	Right cerebrum, occipital lobe, lingual gyrus	19
-22	-75	20	4.3	Left cerebrum, occipital lobe, cuneus	18
-30	-84	-32	4.2	Left cerebellum, posterior lobe, pyramis	
38	-80	-28	5.3	Right cerebellum, posterior lobe, tuber	
4	-80	-34	5.4	Right cerebellum, posterior lobe, uvula	

\* P<0.05 (FDR corrected)

BA, Brodmann area

**Table 4.** 3.0T less than 1.5T WM differences

<b>X</b>	<b>Y</b>	<b>Z</b>	<b>*Z score</b>	<b>Region</b>
-34	10	40	5.3	Left cerebrum, frontal lobe, middle frontal gyrus
36	-4	46	5.2	Right cerebrum frontal lobe, middle frontal gyrus
-12	58	-18	5.1	Left cerebrum, frontal lobe, medial frontal gyrus
-30	42	20	4.3	Left cerebrum, frontal lobe, middle frontal gyrus
26	54	14	4.2	Right cerebrum, frontal lobe, superior frontal gyrus
-52	4	32	5.3	Left cerebrum, frontal lobe, inferior frontal gyrus
-44	34	18	4.8	Left cerebrum, frontal lobe, middle frontal gyrus
24	18	48	3.2	Right cerebrum, frontal lobe, superior frontal gyrus
-16	42	46	5.4	Left cerebrum, frontal lobe, superior frontal gyrus
32	-58	38	3.3	Right cerebrum, parietal lobe, subgyral
52	-44	28	4.6	Right cerebrum, parietal lobe, inferior parietal lobule



-36	-56	34	5.2	Left cerebrum, parietal lobe, angular gyrus
-64	-24	-8	5.1	Left cerebrum, temporal lobe, subgyral
40	-34	-18	3.3	Right cerebrum, temporal lobe, fusiform gyrus
-60	-2	-12	4.6	Left cerebrum, temporal lobe, middle temporal gyrus
58	6	-6	3.2	Right cerebrum, temporal lobe, superior temporal gyrus
-44	-48	-18	3.3	Left cerebrum, temporal lobe, fusiform gyrus
-54	-54	16	4.6	Left cerebrum, temporal lobe, superior temporal gyrus
-8	-14	42	5.2	Left cerebrum, limbic lobe, cingulate gyrus
14	-30	42	5.1	Right cerebrum, limbic lobe, cingulate gyrus
2	-46	14	6.9	Right cerebrum, limbic lobe posterior cingulate gyrus
-26	-56	8	6	Left cerebrum, sublobar, extranuclear
38	-18	2	5.2	Right cerebrum, sublobar, extranuclear
-36	-14	12	5.1	Left cerebrum, sublobar, insula

\* P<0.05 (FDR corrected)

**Table 5.** MD greater than MP GM differences

<b>X</b>	<b>Y</b>	<b>Z</b>	<b>*Z score</b>	<b>Region</b>	<b>BA</b>
22	32	12	7.4	Right cerebrum, limbic lobe, anterior cingulate	32
-62	-46	-10	2.4	Left cerebrum, temporal lobe, middle temporal gyrus	21
-32	28	12	7.3	Left cerebrum, sublobar, insula	13
-54	-44	-32	3.1	Left cerebellum, anterior lobe, culmen	
2	-58	-30	2.4	Right cerebellum, anterior lobe, nodule	
58	-54	-26	2.4	Right cerebellum, posterior lobe, tuber	

\*P<0.05 (FDR corrected)

BA, Brodmann area

**Table 6.** MD greater than MP WM differences

<b>X</b>	<b>Y</b>	<b>Z</b>	<b>*Z score</b>	<b>Region</b>
0	-28	-12	2	Brainstem, midbrain

**Table 7.** MD less than MP GM differences

<b>X</b>	<b>Y</b>	<b>Z</b>	<b>*Z score</b>	<b>Region</b>	<b>BA</b>
-18	28	34	4.7	Left cerebrum, frontal lobe, medial frontal gyrus	6
22	22	32	5.6	Right cerebrum, frontal lobe, cingulate gyrus	32
32	10	26	5.6	Right cerebrum, frontal lobe, inferior frontal gyrus	9
40	-6	26	4.7	Right cerebrum, frontal lobe, precentral gyrus	6
20	-52	24	5	Right cerebrum, parietal lobe, precuneus	31
34	-34	46	4.5	Right cerebrum, parietal lobe, post central gyrus	3
-26	-38	46	4.6	Left cerebrum, parietal lobe, post central gyrus	3
-24	-60	20	4.5	Left cerebrum, limbic lobe, posterior cingulate gyrus	31
-40	-42	2	4.3	Left cerebrum, limbic lobe, parahippocampal gyrus	19
-40	-34	0	4.2	Left cerebrum, temporal lobe, subgyral	
14	12	16	4.5	Right cerebrum, body of caudate nucleus	
36	-46	26	5.9	Right cerebrum, insula	13

\*P<0.05 (FDR corrected)

BA, Brodmann area

**Table 8.** MD less than MP WM differences

<b>X</b>	<b>Y</b>	<b>Z</b>	<b>*Z score</b>	<b>Region</b>
-4	-42	10	5.2	Left cerebrum, corpus callosum
-44	-46	-12	4.9	Left cerebrum, temporal lobe, fusiform gyrus

\*P<0.05 (FDR corrected)

**Table 9.** Unmatched scanning parameters in our study

<b>Scanning parameters</b>	<b>Scanner field strength and sequence</b>		
	<i>1.5 T MPRAGE</i>	<i>3.0 T MPRAGE</i>	<i>3.0 T MDEFT</i>
<i>Magnetization preparation</i>	Non selective inversion recovery	Slice selection inversion recovery	Non selective inversion recovery
<i>Flip angle</i>	15	8	16
<i>FOV/cm</i>	25.6	20	25.6
<i>Matrix size</i>	256 X 224	320 X 320	256 X 240
<i>Voxel size</i>	1.0 X 1.0 X 1.0	0.6 X 0.6 X 1.0	1.0 X 1.0 X 1.0
<i>Bandwidth/Hz/Px</i>	130	190	195

<i>No. of slices</i>	176	192	176
<i>Time of acquisition/min</i>	7.38	12.18	12.51

#### 4.0 Discussion and conclusions

Our results reveal the presence of GM and WM differences in most brain regions between scans obtained at 1.5T and 3.0T as well as between scans obtained with MPRAGE and MDEFT sequences. These differences can be attributed to technical differences like (SNR, CNR and different types of artifacts) between the two comparisons rather than to real anatomical differences as we scanned same subjects in both comparisons. We use VBM technique for comparison[48] as it offers the possibility of examining the whole brain in an objective unbiased manner and has been widely used in health[52, 53] and disease[54-56].

##### 4.1 1.5T versus 3.0T

Our results show GM and WM differences in almost all brain regions including cortical and deep cerebral regions and in cerebellum. These differences were more prevalent with 1.5T less than 3.0T as compared to 1.5T greater than 3.0T contrasts. They may be attributed to a higher SNR and CNR in 3.0T scans and to different artifacts. In theory, the intensity of MR signal is linearly correlated with the strength of the static magnetic field  $B_0$ [4]. So we expect SNR to be doubled in 3.0T albeit we have not measured it. As expected[3], artifacts were higher in 3.0T scans especially motion, chemical shift and susceptibility artifacts. The longer acquisition time as compared to 1.5T may account for the higher observed motion artifacts.

A previous study that used manual tracing to compare hippocampal volumes at 1.5T and 3.0T reveal no inter-scanner differences[7]. Our results partially agree with this study in revealing no differences in the hippocampus although we might be able to reveal such differences if we used a smaller Kernel size such as 4mm [53]. We think that our study may be more reliable and applicable as we used an objective method and examined the whole brain. In neuroscience researches[57] as well as in many psychological and neurological disorders[58, 59], brain changes are often diverse involving more than one brain region. It is therefore extremely important to search for inter-scanner differences over the whole brain including both GM and WM regions rather than restricting search to one cerebral region that might mask the effect of field strength. We address this issue by selecting VBM method to examine the whole brain.

A study that used VBM to examine the effect of hormone replacement therapy on brains of postmenopausal women [60] reported no differences between scans obtained at both 1.5 and 3.0T scanners and, consequently, authors included 26 scans obtained at 1.5T scanner and 17 scans at 3.0T scanner in their voxel-wise comparison. Furthermore, they used scanners of different manufacturers (Milwaukee and Siemens). Based upon results of our study, we disagree with this study as we reported differences even when comparing same subjects and we would expect more differences when comparing different subjects. Likewise, we would expect more differences when using scanners from different manufacturers (all of our scans were obtained from Siemens scanners to decrease bias). We hypothesize that recruiting scans from two different field strengths would introduce several non-biological GM and WM differences. This assumption was kept in mind in nearly all previous VBM studies which prompted them to use scans obtained within the same scanner to ensure validity of their results[61].

3.0T scanners have been increasingly used in research and clinical centers as it was shown that the use of high field strength would enhance image quality thereby increasing number of activated voxels in functional studies[8], improving resolution in MR spectroscopic studies[9] and increasing diagnostic efficacy in clinical studies[10-13]. However, care should be taken when making morphometric comparisons with previous studies committed in 1.5 scanners as some of the differences could be due to inter-scanner technical differences. Equally important not to include scans from different field strengths in a single study such as VBM which though having the advantage of saving time and volunteers, may lead to serious shortcomings.

## 4.2 MPRAGE versus MDEFT

MPRAGE and MDEFT sequences are increasingly used in research to obtain T1 scans with high resolution, CNR and SNR in a relatively short acquisition time. These inversion recovery sequences are preferred to spin echo (SE) sequence for many reasons[4]. It has been shown that using SE may lead to reduced CNR in 3.0T as compared to 1.5T due to prolongation and convergence of T1 times of GM and WM at higher field strengths [41, 62, 63]. In addition, shielding effects induced by eddy currents prevent central parts of the image from being properly excited[64] which results in reduced signal intensity in basal ganglia. Finally, there is enhancement of magnetization transfer effects at higher  $B_0$  that reduce SNR and CNR.

In contrast to inter-scanner differences, there were fewer intra-scanner differences detected when comparing MDEFT versus MPRAGE. We argue that this is because both sequences were acquired within the same scanner (i.e. same field strength) and that the type of field strength is more powerful than the type of pulse sequence in revealing these differences. Similar to inter-scanner parcellation, the differences observed are due to technical factors such as CNR, SNR and artifacts. CNR was reported to be better with MDEFT as compared to MPRAGE sequence and this may be one reason to the observed differences[46, 47]. In addition, artefacts may play a substantial role in provoking these differences. We might exclude the effect of motion artefacts from our hypotheses as acquisition time was almost identical in both sequences although slightly higher with the MDEFT one. We hypothesize that susceptibility artefacts were the major type of artefacts responsible to the observed differences. Susceptibility artefacts occur as the result of microscopic gradients or variations in the magnetic field strength that occurs near the interfaces of substance of different magnetic susceptibility. These gradients cause dephasing of spins and frequency shifts of the surrounding tissues. The net result is bright and dark areas with spatial distortion of surrounding anatomy[65]. These artefacts are worst with gradient echo sequences such as the MPRAGE and MDEFT. However, MDEFT was hypothesized to be less sensitive to susceptibility artefacts[3] which we clearly observed in our study (Figure 6).

These results show that even when selecting scans from the same scanner, non-biological GM and WM density differences may be introduced if different pulse sequences were used to obtain them thereby pointing out to the need for stabilizing the confounding effect of pulse sequence in future MR studies.

In conclusion, the study reveals in an objective, unbiased, whole brain analysis that several GM and WM differences do exist between scans obtained at different field strengths and with different pulse sequences. As we compared scans of same subjects in both conditions, differences observed are not real biological differences rather they result from technical variations such as SNR, CNR and artifacts. In this context, we propose that combining scans obtained at different field strengths or with different pulse sequences in MR studies is not without the bias of introducing technical variabilities that might adversely affect validity and reliability of these studies. The same precaution should be considered when comparing results obtained from studies using different field strengths and/or different pulse sequences.

## 4.3 Limitations of the study

Although our findings of GM and WM differences could be attributed to using different field strengths and different pulse sequences, they may also be due to other parameters that were not unified across comparisons and so may partly account for the final differences observed. (Table 9) shows these parameters which include magnetization preparation, voxel size, and bandwidth, number of slices, matrix size and field of view. These unmatched parameters might adversely affected our VBM comparison leading to GM and WM differences not merely attributed to field strength or pulse sequence type.

## Acknowledgments

I would like to thank Dr. Vanessa Sluming who supervised this work and provided invaluable suggestions and instructions. I thank Professor Dr. Ralf Deichmann (University Hospital ZNN, Brain Imaging Centre, Frankfurt/Main Germany) for his authorization to implement the MDEFT sequence in the 3.0 T scanner at MARIARC and for his invaluable information and advices regarding MDEFT and MPRAGE sequences comparison. I also thank Dr. Laura Parkes

(MARIARC, University of Liverpool) for her helpful suggestions and Dr Guillaume Flandin, Wellcome Trust Centre for Neuroimaging, University College London for her kind responses regarding SPM5 and VBM data analysis. Finally, I would like to thank Mrs Valerie Adams (research radiographer at MARIARC, University of Liverpool) for her great help in data acquisition and subjects' preparation.

## References

- Frayne, R., et al., *Magnetic resonance imaging at 3.0 Tesla: challenges and advantages in clinical neurological imaging*. Invest Radiol, 2003. **38**(7): pp. 385-402.
- Schick, F., *Whole-body MRI at high field: technical limits and clinical potential*. Eur Radiol, 2005. **15**(5): pp. 946-59.
- Norris, D.G., *High field human imaging*. J Magn Reson Imaging, 2003. **18**(5): pp. 519-29.
- Schmitz, B.L., et al., *Advantages and pitfalls in 3T MR brain imaging: a pictorial review*. AJNR Am J Neuroradiol, 2005. **26**(9): pp. 2229-37.
- Ogawa, S., et al., *Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields*. Magn Reson Med, 1990. **14**(1): pp. 68-78.
- Ogawa, S., et al., *Brain magnetic resonance imaging with contrast dependent on blood oxygenation*. Proc Natl Acad Sci U S A, 1990. **87**(24): pp. 9868-72.
- Briellmann, R.S., A. Syngeniotis, and G.D. Jackson, *Comparison of hippocampal volumetry at 1.5 tesla and at 3 tesla*. Epilepsia, 2001. **42**(8): pp. 1021-4.
- Tieleman, A., et al., *Comparison between functional magnetic resonance imaging at 1.5 and 3 Tesla: effect of increased field strength on 4 paradigms used during presurgical work-up*. Invest Radiol, 2007. **42**(2): pp. 130-8.
- Barker, P.B., D.O. Hearshen, and M.D. Boska, *Single-voxel proton MRS of the human brain at 1.5T and 3.0T*. Magn Reson Med, 2001. **45**(5): pp. 765-9.
- Nielsen, K., et al., *Magnetic resonance imaging at 3.0 tesla detects more lesions in acute optic neuritis than at 1.5 tesla*. Invest Radiol, 2006. **41**(2): pp. 76-82.
- Bachmann, R., et al., *FLAIR imaging for multiple sclerosis: a comparative MR study at 1.5 and 3.0 Tesla*. Eur Radiol, 2006. **16**(4): pp. 915-21.
- Nobauer-Huhmann, I.M., et al., *Magnetic resonance imaging contrast enhancement of brain tumors at 3 tesla versus 1.5 tesla*. Invest Radiol, 2002. **37**(3): pp. 114-9.
- Ba-Ssalamah, A., et al., *Effect of contrast dose and field strength in the magnetic resonance detection of brain metastases*. Invest Radiol, 2003. **38**(7): pp. 415-22.
- Fischl, B., et al., *Sequence-independent segmentation of magnetic resonance images*. Neuroimage, 2004. **23** Suppl 1: pp. S69-84.
- Mugler, J.P., 3rd and J.R. Brookeman, *Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE)*. Magn Reson Med, 1990. **15**(1): pp. 152-7.
- Mugler, J.P., 3rd and J.R. Brookeman, *Rapid three-dimensional T1-weighted MR imaging with the MP-RAGE sequence*. J Magn Reson Imaging, 1991. **1**(5): pp. 561-7.
- Ugurbil, K., et al., *Imaging at high magnetic fields: initial experiences at 4 T*. Magn Reson Q, 1993. **9**(4): pp. 259-77.
- Lee, J.H., et al., *High contrast and fast three-dimensional magnetic resonance imaging at high fields*. Magn Reson Med, 1995. **34**(3): pp. 308-12.
- Epstein, F.H., J.P. Mugler, 3rd, and J.R. Brookeman, *Optimization of parameter values for complex pulse sequences by simulated annealing: application to 3D MP-RAGE imaging of the brain*. Magn Reson Med, 1994. **31**(2): pp. 164-77.
- Deichmann, R., et al., *Optimization of 3-D MP-RAGE sequences for structural brain imaging*. Neuroimage, 2000. **12**(1): pp. 112-27.
- Williams, L.A., et al., *Optimization of 3D MP-RAGE for neonatal brain imaging at 3.0 T*. Magn Reson Imaging, 2007. **25**(8): pp. 1162-70.

- Norris, D.G., et al., *MDEFT imaging of the human brain at 8 T*. *MAGMA*, 1999. **9**(1-2): pp. 92-6.
- Deichmann, R., C. Schwarzbauer, and R. Turner, *Optimisation of the 3D MDEFT sequence for anatomical brain imaging: technical implications at 1.5 and 3 T*. *Neuroimage*, 2004. **21**(2): pp. 757-67.
- Frahm, J., A. Haase, and D. Matthaei, *Rapid three-dimensional MR imaging using the FLASH technique*. *J Comput Assist Tomogr*, 1986. **10**(2): pp. 363-8.
- Decety, J., et al., *The neural bases of cooperation and competition: an fMRI investigation*. *Neuroimage*, 2004. **23**(2): pp. 744-51.
- Krainik, A., et al., *Regional impairment of cerebrovascular reactivity and BOLD signal in adults after stroke*. *Stroke*, 2005. **36**(6): pp. 1146-52.
- Kruggel, F., *MRI-based volumetry of head compartments: normative values of healthy adults*. *Neuroimage*, 2006. **30**(1): pp. 1-11.
- Peinemann, A., et al., *Executive dysfunction in early stages of Huntington's disease is associated with striatal and insular atrophy: a neuropsychological and voxel-based morphometric study*. *J Neurol Sci*, 2005. **239**(1): pp. 11-9.
- Salgado-Pineda, P., et al., *Brain T1 intensity changes after levodopa administration in healthy subjects: a voxel-based morphometry study*. *Br J Clin Pharmacol*, 2006. **62**(5): pp. 546-51.
- Good, C.D., et al., *Cerebral asymmetry and the effects of sex and handedness on brain structure: a voxel-based morphometric analysis of 465 normal adult human brains*. *Neuroimage*, 2001. **14**(3): pp. 685-700.
- Maguire, E.A., K. Woollett, and H.J. Spiers, *London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis*. *Hippocampus*, 2006. **16**(12): pp. 1091-101.
- Gallinat, J., et al., *Smoking and structural brain deficits: a volumetric MR investigation*. *Eur J Neurosci*, 2006. **24**(6): pp. 1744-50.
- Lee, H., et al., *Anatomical traces of vocabulary acquisition in the adolescent brain*. *J Neurosci*, 2007. **27**(5): pp. 1184-9.
- Brant-Zawadzki, M.N., et al., *Three-dimensional MR imaging and display of intracranial disease: improvements with the MP-RAGE sequence and gadolinium*. *J Magn Reson Imaging*, 1993. **3**(4): pp. 656-62.
- Filippi, M., et al., *A high-resolution three-dimensional T1-weighted gradient echo sequence improves the detection of disease activity in multiple sclerosis*. *Ann Neurol*, 1996. **40**(6): pp. 901-7.
- Fellner, F., et al., *A T1-weighted rapid three-dimensional gradient-echo technique (MP-RAGE) in preoperative MRI of intracranial tumours*. *Neuroradiology*, 1996. **38**(3): pp. 199-206.
- Held, P., et al., *MRI of inner ear and facial nerve pathology using 3D MP-RAGE and 3D CISS sequences*. *Br J Radiol*, 1997. **70**(834): pp. 558-66.
- Held, P., et al., *MRI of inner ear anatomy using 3D MP-RAGE and 3D CISS sequences*. *Br J Radiol*, 1997. **70**(833): pp. 465-72.
- Held, P., et al., *Three-dimensional MP-RAGE--an alternative to conventional three-dimensional FLASH sequences for the diagnosis of viscerocranial tumours?* *Br J Radiol*, 1995. **68**(816): pp. 1316-24.
- Julin, P., et al., *Reliability of interactive three-dimensional brain volumetry using MP-RAGE magnetic resonance imaging*. *Psychiatry Res*, 1997. **76**(1): p. 41-9.
- Wansapura, J.P., et al., *NMR relaxation times in the human brain at 3.0 tesla*. *J Magn Reson Imaging*, 1999. **9**(4): pp. 531-8.
- Fischbach, F., et al., *Efficacy of contrast medium use for neuroimaging at 3.0 T: utility of IR-FSE compared to other T1-weighted pulse sequences*. *J Comput Assist Tomogr*, 2005. **29**(4): pp. 499-505.
- Runge, V.M., et al., *Clinical comparison of three-dimensional MP-RAGE and FLASH techniques for MR imaging of the head*. *J Magn Reson Imaging*, 1991. **1**(4): pp. 493-500.

- Mugler, J.P., 3rd, F.H. Epstein, and J.R. Brookeman, *Shaping the signal response during the approach to steady state in three-dimensional magnetization-prepared rapid gradient-echo imaging using variable flip angles*. Magn Reson Med, 1992. **28**(2): pp. 165-85.
- Thomas, D.L., et al., *3D MDEFT imaging of the human brain at 4.7 T with reduced sensitivity to radiofrequency inhomogeneity*. Magn Reson Med, 2005. **53**(6): pp 1452-8.
- Tardif, C.L. and G.B. Pike, *Comparison of T1-weighted 3D High-resolution anatomical sequences for the brain at 3 Tesla: FLASH, MP-RAGE and MDEFT*, in *Proc. Intl. Soc. MAG. Reson. Med.* 2007: Berlin. pp. 210.
- Furlong, C., et al., *Comparison of MRI sequences for surface area quantification in post mortem brains*, in *BC ISMRM 2005*: Oxford.
- Ashburner, J. and K.J. Friston, *Voxel-based morphometry--the methods*. Neuroimage, 2000. **11**(6 Pt 1): pp. 805-21.
- Ashburner, J. and K.J. Friston, *Unified segmentation*. Neuroimage, 2005. **26**(3): pp. 839-51.
- Talairach, J. and P. Tournoux, *Coplanar stereotaxic atlas of the human brain: three-dimensional proportional system: an approach to cerebral imaging*. 1988, New York: Thieme.
- Brodmann, K., *Vergleichende lokalizationslehre der grosshirnrinde*. Johann Ambrosius Barth, 1909: pp. 130-139.
- Sluming, V., et al., *Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians*. Neuroimage, 2002. **17**(3): pp. 1613-22.
- Maguire, E.A., et al., *Navigation-related structural change in the hippocampi of taxi drivers*. Proc Natl Acad Sci U S A, 2000. **97**(8): pp. 4398-403.
- Li, S., et al., *Regional white matter decreases in Alzheimer's disease using optimized voxel-based morphometry*. Acta Radiol, 2008. **49**(1): pp. 84-90.
- Garcia-Marti, G., et al., *Schizophrenia with auditory hallucinations: A voxel-based morphometry study*. Prog Neuropsychopharmacol Biol Psychiatry, 2008. **32**(1): pp. 72-80.
- Kim, J.H., et al., *Regional grey matter abnormalities in juvenile myoclonic epilepsy: a voxel-based morphometry study*. Neuroimage, 2007. **37**(4): pp. 1132-7.
- Gaser, C. and G. Schlaug, *Brain structures differ between musicians and non-musicians*. J Neurosci, 2003. **23**(27): pp. 9240-5.
- Kubicki, M., et al., *Voxel-based morphometric analysis of gray matter in first episode schizophrenia*. Neuroimage, 2002. **17**(4): pp. 1711-9.
- Frisoni, G.B., et al., *Detection of grey matter loss in mild Alzheimer's disease with voxel based morphometry*. J Neurol Neurosurg Psychiatry, 2002. **73**(6): pp. 657-64.
- Erickson, K.I., et al., *Selective sparing of brain tissue in postmenopausal women receiving hormone replacement therapy*. Neurobiol Aging, 2005. **26**(8): pp. 1205-13.
- Mechelli, A., et al., *Voxel-Based Morphometry of the Human Brain: Methods and Applications*. Current Medical Imaging Reviews, 2005, **1**: pp. 1-9.
- Ethofer, T., et al., *Comparison of longitudinal metabolite relaxation times in different regions of the human brain at 1.5 and 3 Tesla*. Magn Reson Med, 2003. **50**(6): pp. 1296-301.
- Michaeli, S., et al., *Proton T2 relaxation study of water, N-acetylaspartate, and creatine in human brain using Hahn and Carr-Purcell spin echoes at 4T and 7T*. Magn Reson Med, 2002. **47**(4): pp. 629-33.
- Hoult, D.I. and D. Phil, *Sensitivity and power deposition in a high-field imaging experiment*. J Magn Reson Imaging, 2000. **12**(1): pp. 46-67.
- Mirowitz, S.A., *MR imaging artifacts. Challenges and solutions*. Magn Reson Imaging Clin N Am, 1999. **7**(4): pp. 717-32.

# Examining and Improving the Limitations of the Gazis-Herman-Rothery Car-following Model

Hamid Al-Jameel, School of the Built Environment, University of Salford

## Abstract

Simulation models enable one to represent the most complex situations, and are therefore effective tools for solving traffic problems. Different car-following models have been used to simulate traffic movements. The car-following model known as Gazis-Herman-Rothery (GHR) is one of the earliest models; it has been used since the late 1950`s and up to this time. In this model, the acceleration of follower is based on the spacing and relative speeds between two vehicles (the follower and the leader).

This model has many limitations. For example, there is no response (or zero acceleration or deceleration) for the follower when the relative speed between the leader and follower is equal to zero for any relative spacing between the two vehicles. Another limitation involves the effect of the leading vehicle on its follower even when the distance between them is great. This effect hinders the following vehicle from reaching its desired speed, thereby reflecting unrealistic behaviour.

This paper tries to overcome the effect of large spacing due to car-following by using a spacing threshold to transfer the car-following regime to a free-flow regime.

Finally, sensitivity analysis has been conducted to select the values for parameters in cases of acceleration and deceleration. The modified model has been validated by using real data. The parameters of the model have been selected, and suitable spacing threshold values have been used in calibrating the model.

**Keywords:** Traffic micro-simulation, car-following model, GHR model

## 1. Introduction

Car following rules can be considered as the basic unit for building any traffic-simulation model since they describe the interaction between vehicles travelling at close distances (Brackstone and MacDonald, 1999). During the last five decades, several car-following models were used to interpret the influence of the leading vehicle on the following vehicle. This effect by the leading vehicle can be expressed in the following relationship:

$$\text{Response} = \text{Sensitivity} \times \text{Stimulus} \text{ (Equation 1)}$$

where response represents the acceleration or deceleration of the follower; sensitivity represents a constant or a function of the follower's speed and spacing between the follower and its leader; and stimulus represents the difference in speed between the follower and its leading vehicle

The Gazis-Herman-Rothery model (GHR) is an example of the response-stimulus relationship.

## 2. The GHR model

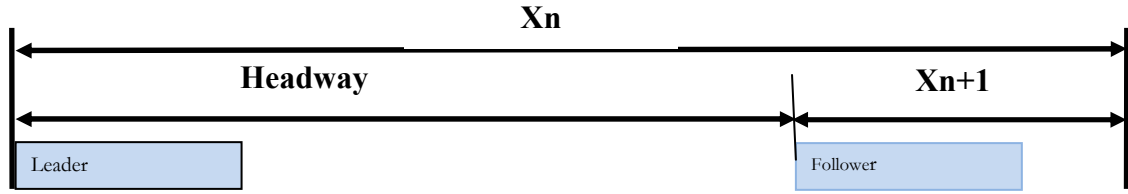
The GHR model is considered one of the earliest car-following models. This model was developed by the General Motors Research Laboratory in Detroit in 1958 (Brackstone and McDonald, 1999). The model has been improved by different researchers over the last five decades.

According to Brackstone and McDonald (1999), this model is represented by the following mathematical equation (see also Figure 1):

$$a_{n+1}(t+T) = c (V_{n+1}(t+T))^m [V_n(t) - V_{n+1}(t)] / [X_n(t) - X_{n+1}(t)]^L \text{ (Equation 2)}$$

where  $a_{n+1}(t+T)$  = acceleration/deceleration of the follower;  $X_n(t)$ ,  $X_{n+1}(t)$  = position of leader and follower, respectively;  $V_n(t)$ ,  $V_{n+1}(t)$  = speed of the leader and follower, respectively; and  $L$ ,  $c$ , and  $m$  represent constant parameters.





**Figure 1.** A sketch illustrating the terms used for the relative positions of the follower and its leader.

According to Equation 1,  $a_{n+1}(t+T)$  represents response. The difference in speed represents stimulus, whereas the speed of the follower and spacing represent sensitivity.

### 2.1. Development of the GHR model

For several decades, researchers have significant interest in the GHR model. Table 1 shows a summary of the development of this model. Moreover, the most important parameters that have been found by different researchers are shown in Table 2.

### 2.2. Limitations of the GHR Model

The GHR model suffers from several limitations. Some of them have been resolved during its long history of development (see Table 1). Despite the improvements of the GHR model, it still needs more refinement to better represent real-life traffic behaviour. The important limitations of the GHR model can be summarised as follows:

- a. The follower reacts to any small changes in the relative speed of its leader (Olstam and Tapani, 2004).
- b. The follower is affected by its leader even if the distance between them is significant. To overcome this problem, a deterministic space threshold has been used as a separation between car-following and free-flow regimes. If the spacing is more than this threshold, the follower is considered to be unaffected by its leader. In this case, the follower will drive to attain his desired speed. Herman et al. (1959) assumed a value of 61 metres as a deterministic space threshold. Aycin (2001) used 250 feet ( $\approx 75$  metres) as a deterministic threshold, along with another limitation on the value of deceleration to move from the car-following to a free-flow state. Toledo (2003) used a similar value of 76 metres.
- c. There is no obvious connection between driver's behaviour and the parameters  $c$ ,  $m$  and  $L$  used in the GHR model (Gipps, 1981).

**Table 1.** Most important historical development for the GHR model.

Researcher(s)	Findings	Mathematical expression
<b>Chandler et al. (1958)</b>	Acceleration/deceleration depends on relative speed only; it does not depend on the spacing between the follower and its leader.	$a_{n+1}(t+T) = c [V_n(t) - V_{n+1}(t)]$
<b>Herman et al. (1959)</b>	A new factor has been added: the spacing between the following and leading vehicles.	$a_{n+1}(t+T) = \frac{c [V_n(t) - V_{n+1}(t)]}{[X_n(t) - X_{n+1}(t)]}$
<b>Herman et al. (1959) (see also Brackstone and McDonald)</b>	Calibrated the preceding model and found good results with $r^2=0.8-0.9$ .	$a_{n+1}(t+T) = \frac{c [V_n(t) - V_{n+1}(t)]}{[X_n(t) - X_{n+1}(t)]}$

<b>[1999]</b>		
<b>Eddie (1961)</b>	Added follower's speed as a new factor.	$a_{n+1}(t+T) = \frac{C (V_{n+1}(t+T)) [V_n(t) - V_{n+1}(t)]}{[X_n(t) - X_{n+1}(t)]}$
<b>Gazis et al. (1961)</b>	Used different values of speed and spacing in terms of m and L.	$a_{n+1}(t+T) = \frac{C (V_{n+1}(t+T))^m [V_n(t) - V_{n+1}(t)]}{[X_n(t) - X_{n+1}(t)]^L}$
<b>May and Keller (1967)</b>	Calibrated the parameters of the sensitivity factors introduced by Eddie in 1961.	Used m=1 and L=3
<b>Aron (1988) (See also Brackstone and McDonald [1999])</b>	Classified driver's response into deceleration, constant speed, and acceleration.	Different values of m and L in different cases; acceleration, constant speed and deceleration.
<b>Ozaki (1993) (See also Brackstone and McDonald [1999])</b>	Also used different sensitivity values.	Different values of m and L in different cases.

**Table 2.** Most reliable estimates of the parameters m and L within the GHR model according to Brackstone and McDonald (1999).

Source	m	L
<b>Chandler et al. (1958)</b>	0	0
<b>Herman et al. (1959)</b>	0	1
<b>Hoefs (1972) (dcn no brk / dcn brk / acn)</b>	1.5/.2/.6	.9/.9/3.2
<b>Treiterer and Myers (1974) (dcn / acn)</b>	.7/.2	2.5/1.6
<b>Ozaki (1993) (dcn / acn)</b>	.9/-.2	1/.2

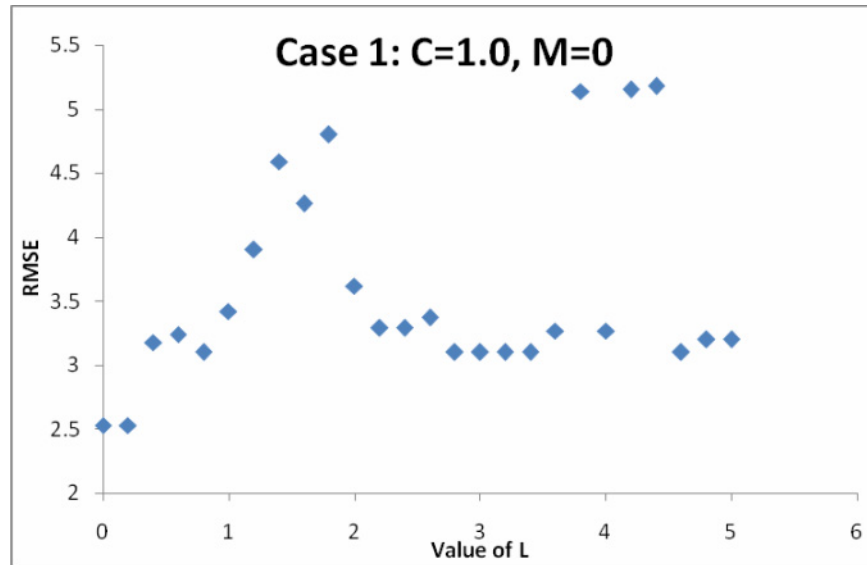
Note: dcn/acn: deceleration/acceleration and brk/no brk: deceleration with and without the use of brakes.

### 3. The Model/Program

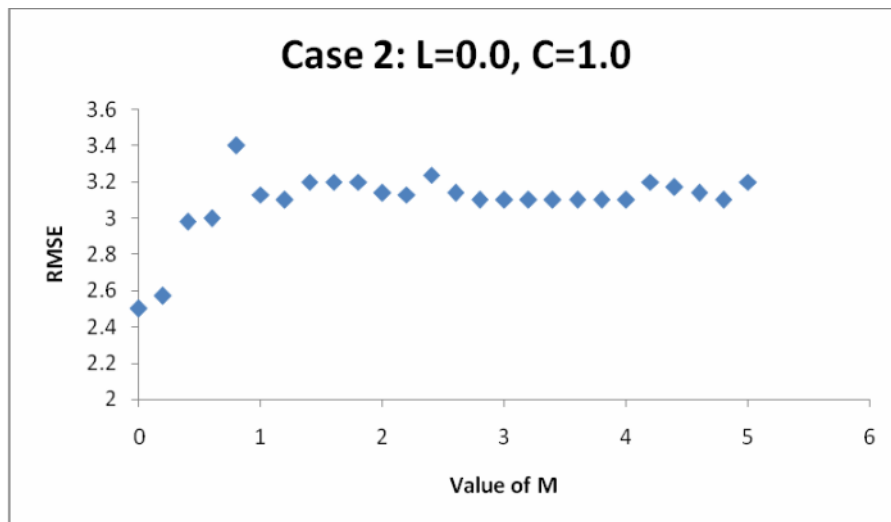
In this paper, a computer model was developed based on the mathematical equations representing the GHR model. The model was built by using Visual FORTRAN Language (6.5). The real data used to calibrate the model consists of two tests from California, USA. The data was collected during daylight by using video cameras and radars to measure the speed and spacing for the two vehicles—that is, the leader and the follower (Saechan, 2009). In the computer program, actual field data for the leader has been used in order to see how the program mimics reality by comparing data for the follower with that obtained from the field. The real data (filed data) consists of two parts: one for speed and another for space headway.

The statistical test used in this calibration is the Root Mean Square Error (RMSE). This statistical test is used as an index to show any differences between the computed and observed data. The results take a value equal or greater to zero. If the result is zero, it is considered an ideal value, and vice versa.

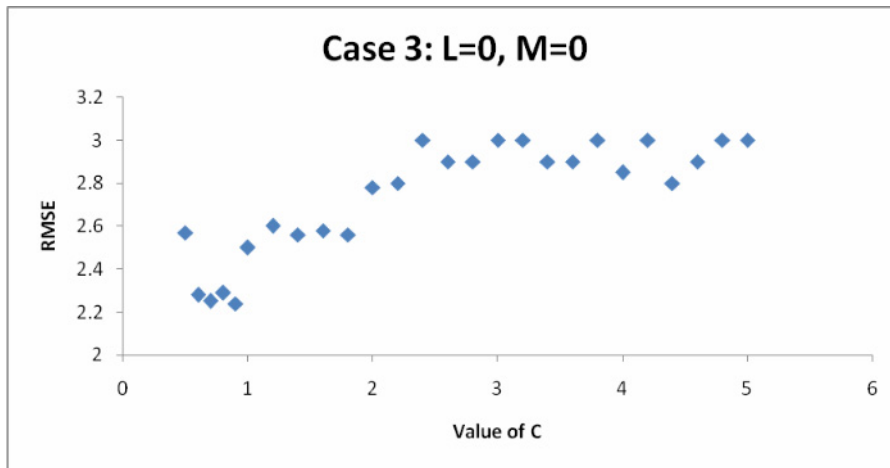
A calibration process has been conducted to select the parameters  $L$ ,  $m$ , and  $c$ . These parameters represent the constants of the GHR model as shown in Equation 2. Different values have been used, ranging from 0 to 5, 0 to 3, and 0 to 10 (with an increment of 0.2) for  $L$ ,  $m$ , and  $c$ , respectively. Several iterations have been carried out in the model to get the best values for these constants. In these iterations, both symmetrical and unsymmetrical behaviour have been tested. For symmetrical behaviour, the same value for the model parameters have been used for acceleration and deceleration cases, whereas for the unsymmetrical behaviour, a certain value has been used for the parameter in case of acceleration and a different one for deceleration. The results of the calculations have been presented in Figures 2–7 (Cases 1–6). For each of these six cases, one parameter is changed and the others remain constant in order to investigate the relationship between RMSE and each variable separately. Depending on the data from both field tests, the best combinations of these parameters are as shown in Table 3.



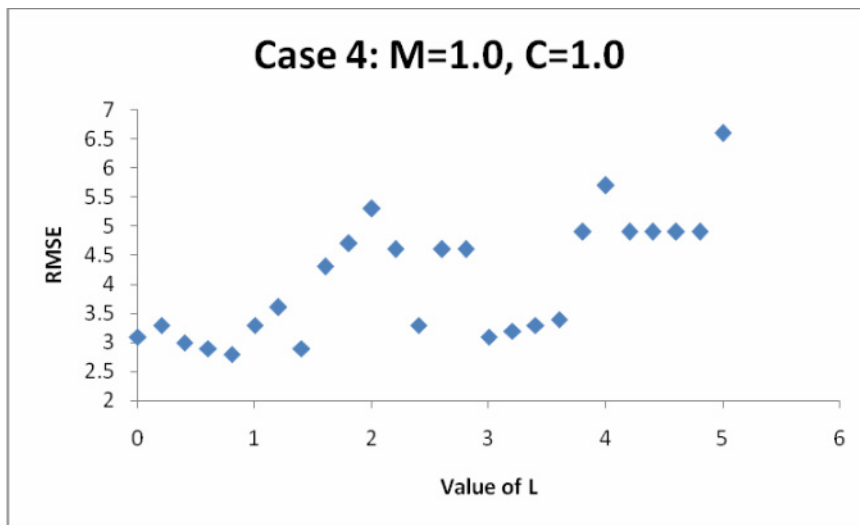
**Figure 2.** RMSE vs.  $L$  for  $c=1$ ,  $m=0$ .



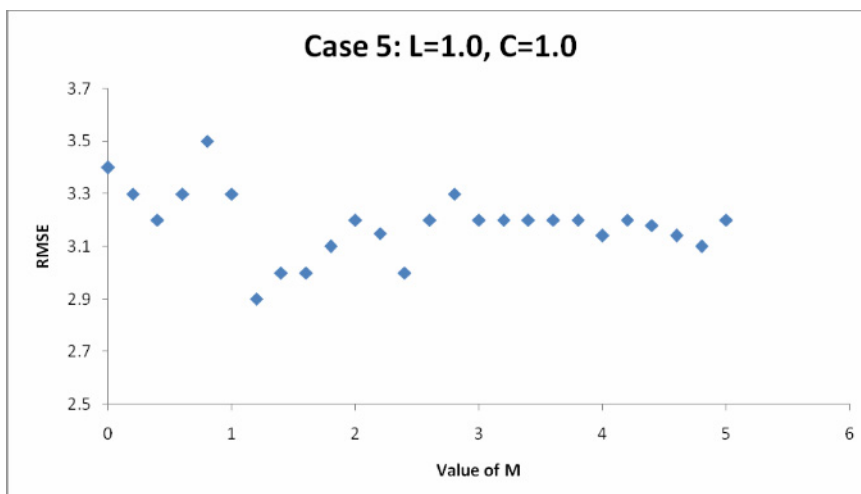
**Figure 3.** RMSE vs.  $m$  for  $L=0$ ,  $c=1$ .



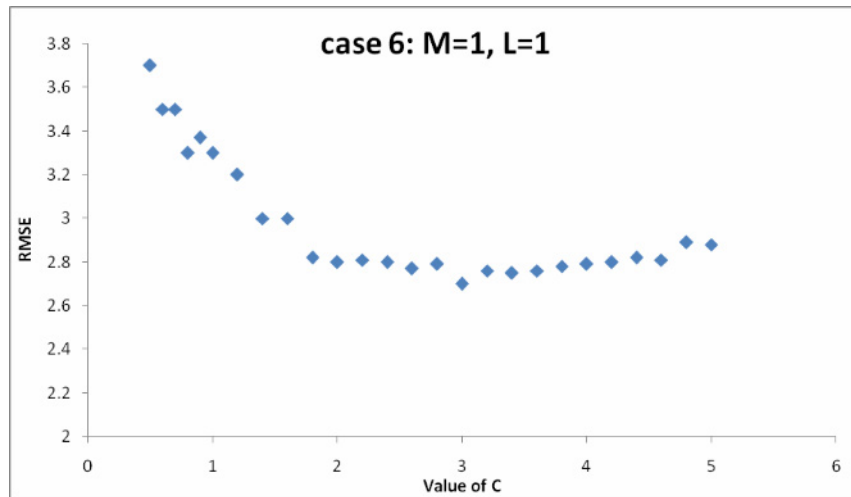
**Figure 4.** RMSE vs.  $c$  for  $L=0$ ,  $m=0$ .



**Figure 5.** RMSE vs.  $L$  for  $m=1$ ,  $c=1$ .



**Figure 6.** RMSE vs.  $m$  for  $L=1$ ,  $c=1$ .



**Figure 7.** RMSE vs. c for m=1, L=1.

**Table 3.** Results of the best combination of the calibrated GHR parameters.

Combination of the parameters	M	L	c	RMSE
❖ Acceleration:	1.0	1.0	1.0	0.78
❖ Deceleration:	1.0	1.3	1.2	
❖ Acceleration:	1.0	1.0	1.0	0.79
❖ Deceleration:	1.1	1.4	1.2	
❖ Acceleration:	1.0	1.1	1.0	0.75
❖ Deceleration:	1.1	1.4	1.0	
❖ Acceleration:	1.0	1.1	1.2	0.55
❖ Deceleration:	1.1	1.4	1.2	

#### 4. Calibration

The calibration has been carried out for two tests at different spacing thresholds. The real data has been used in the program to find the optimum value of spacing threshold. Each test is discussed separately in the following sections.

##### 4.1. Calibration—Test 1

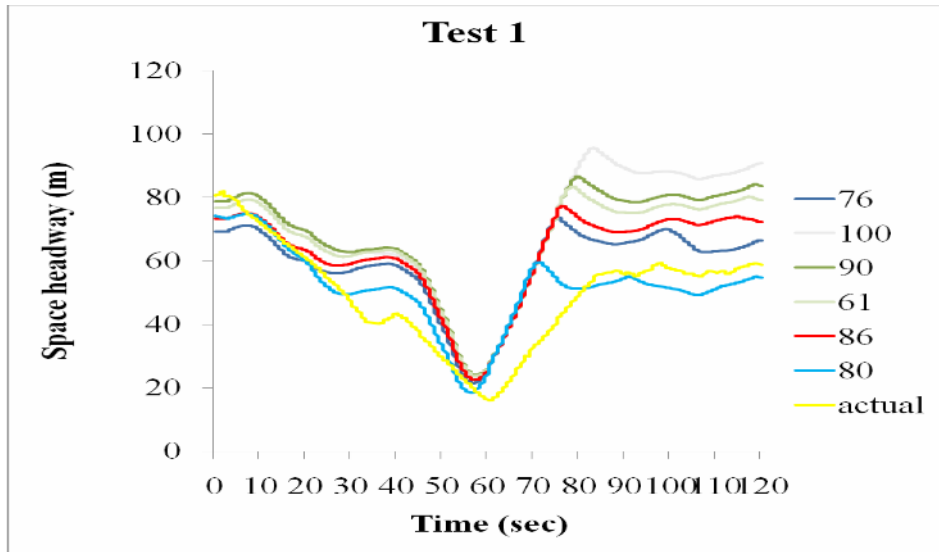
In this section, the selection of the best value of spacing threshold will be discussed. Test 1 is one of two tests that were carried out in the USA, as discussed above. This test consists of having two vehicles—a leader and a follower—travelling over a period of 120 seconds. In addition, the test consists of two parts of real data for speed and space headway. In each part, there is a need to know the optimum value for a spacing threshold from different selected values of these spacing thresholds for the actual speed and space headway data. Beside this, different values of spacing threshold, which were used here, have been taken from previous studies such as Herman et al. (1959) and Toledo (2003).

Figure 8 shows the different values of spacing thresholds for simulated vehicle behaviour against the space headway for actual vehicle behaviour for a period of 120 seconds. It indicates that 80 metres gave the best results in terms of closeness to the actual field data. This is shown separately in Figure 9.

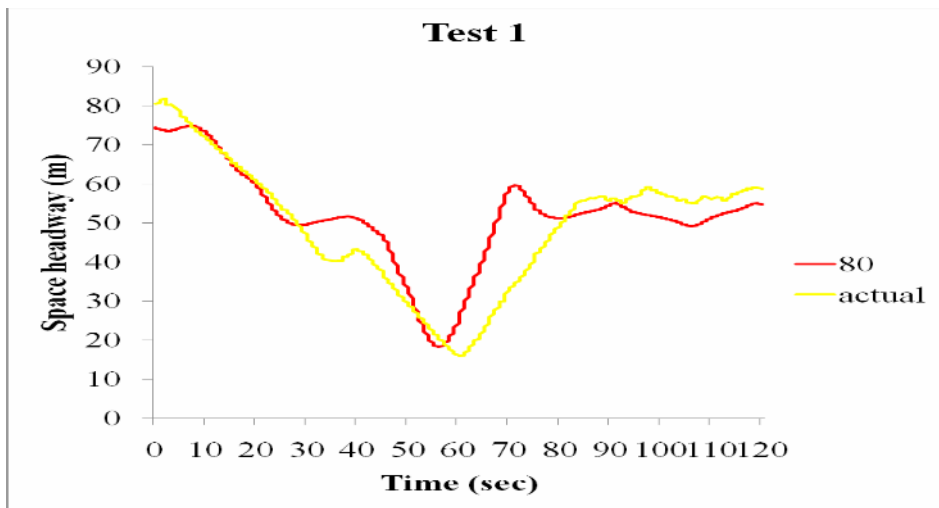
Also, when testing for RMSE, the 80 m spacing threshold represents the optimum value, as shown in Figure 10.

Figure 11 shows how different simulated spacing thresholds vary against actual ones in terms of real speed data. Figures 11–13 indicate that the 80 m spacing threshold is the best value.

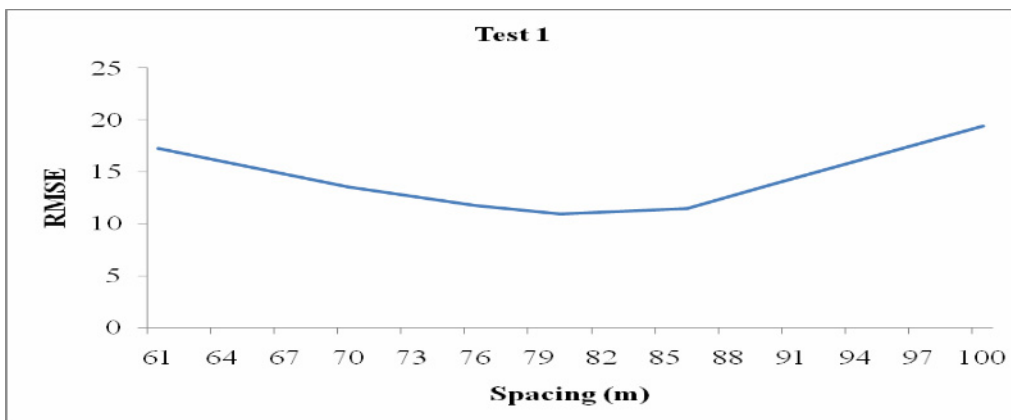
Figure 13 indicates that the 80 m value represents the minimum value of RMSE (i.e., the optimum value).



**Figure 8.** Different values of spacing threshold compared with actual data for space headways.

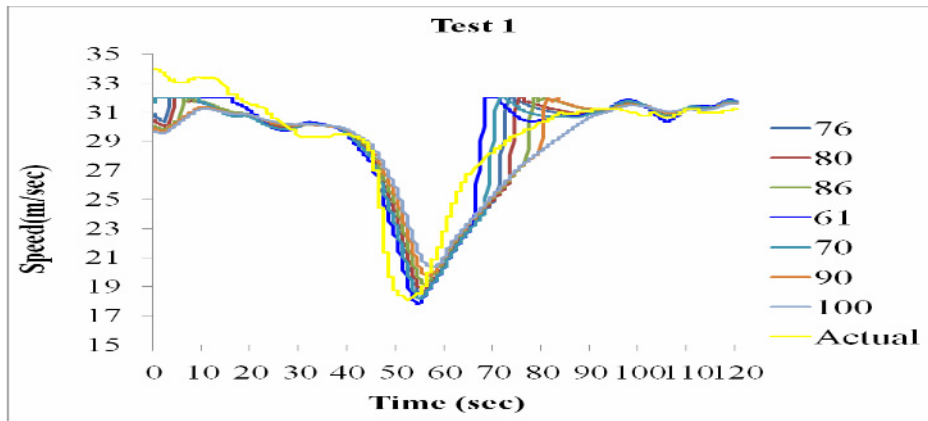


**Figure 9.** Actual field data versus 80 metre threshold value for space headways.

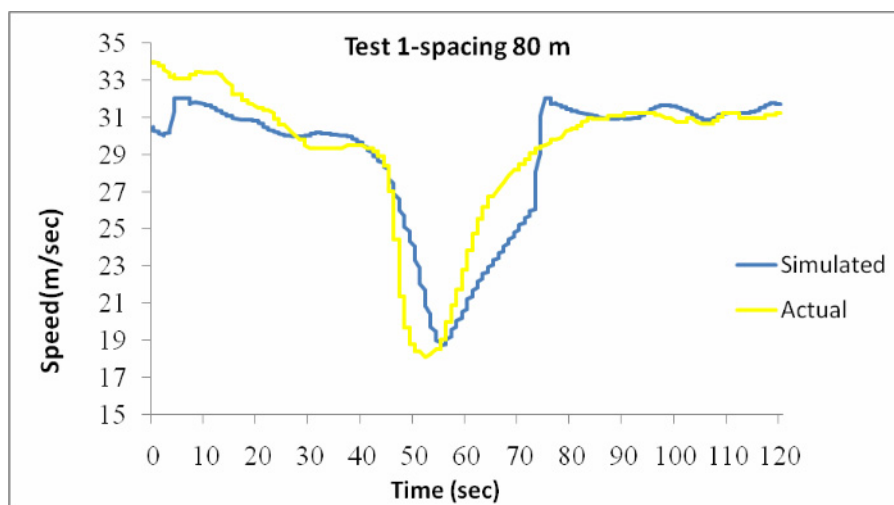


**Figure 10.** Different values of spacing threshold against RMSE values for space headways.

In summary, the optimum value of a spacing threshold for both speed and space headway when compared with real data is 80 m.



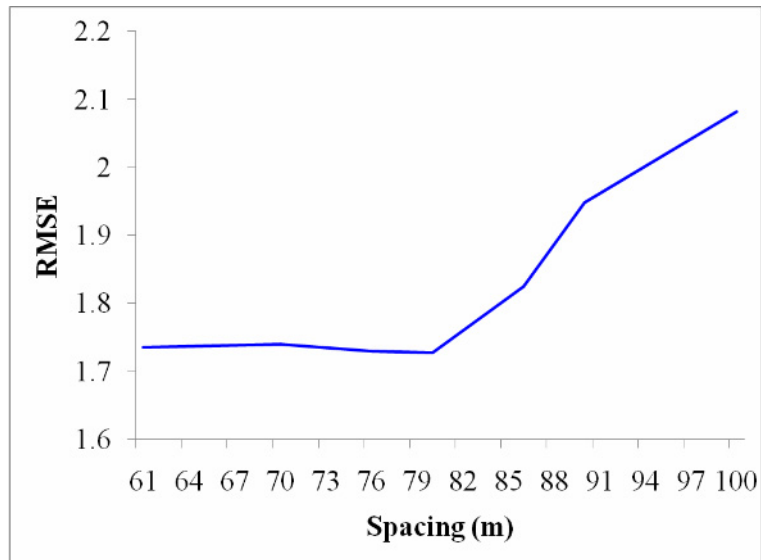
**Figure 11.** Different values of spacing threshold compared with actual data for speeds.



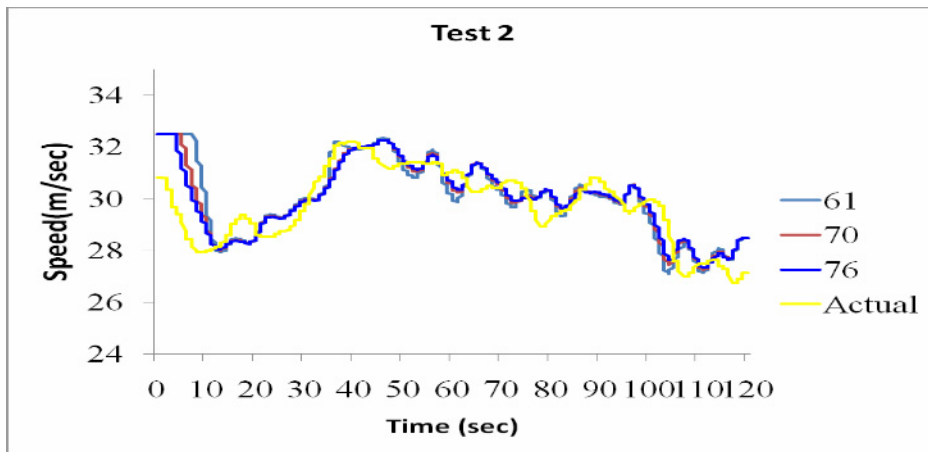
**Figure 12.** Simulated values of speed compared with actual data.

#### 4.2. Calibration—Test 2

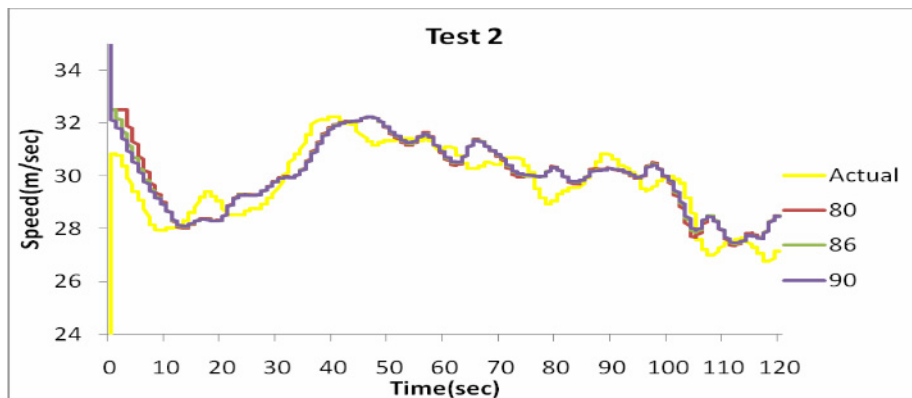
Again, this test consists of two parts of real data: speed and space headway. To obtain the optimum value for the spacing threshold, different values of spacing threshold have been tested against actual data. Figures 14 and 15 represent different values of spacing threshold versus real speed data. Figure 16 shows actual field data versus a value of 80 m. To determine the optimum value, the minimum RMSE is used. Figure 17 shows that a spacing threshold of 86 m is the optimum.



**Figure 13.** Spacing thresholds against RMSE values.

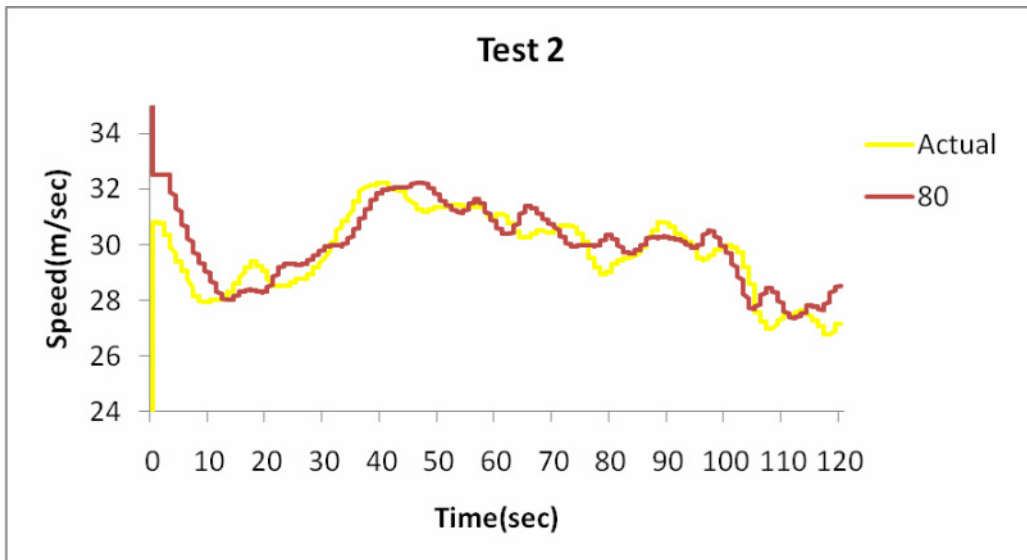


**Figure 14.** Different values of spacing threshold compared with actual data for speeds.

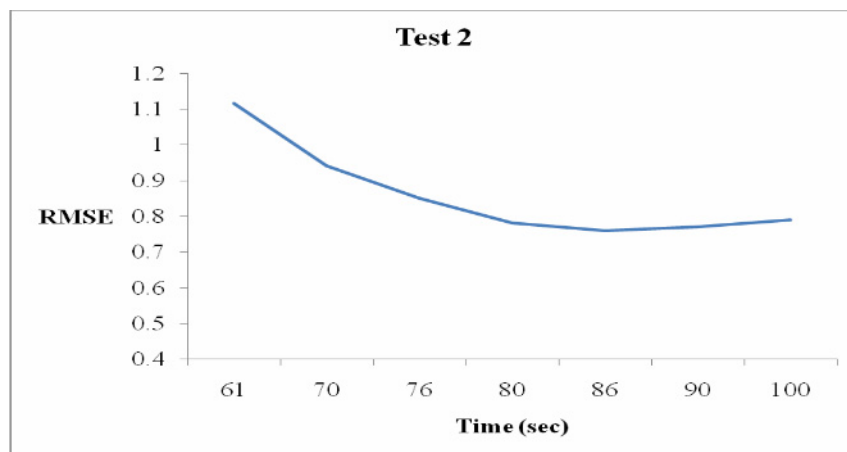


**Figure 15.** Different values of spacing threshold compared with actual data for speeds.





**Figure 16.** Values of spacing threshold at 80 m compared with actual data.

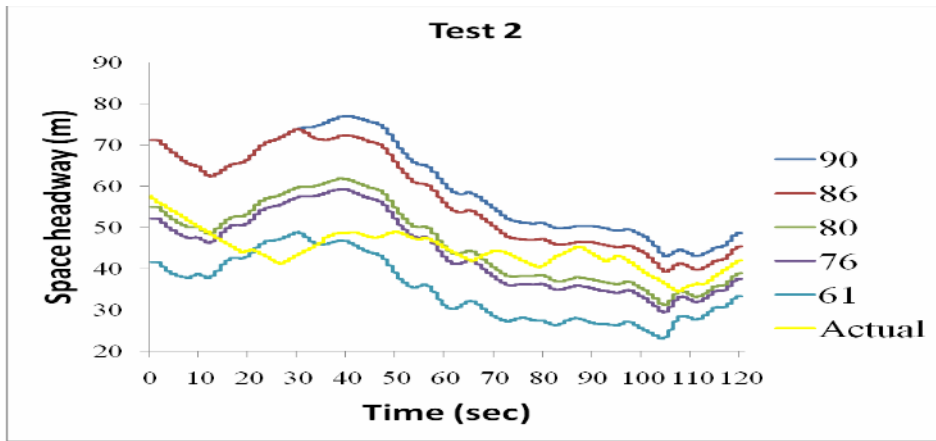


**Figure 17.** Different values of spacing threshold against RMSE values.

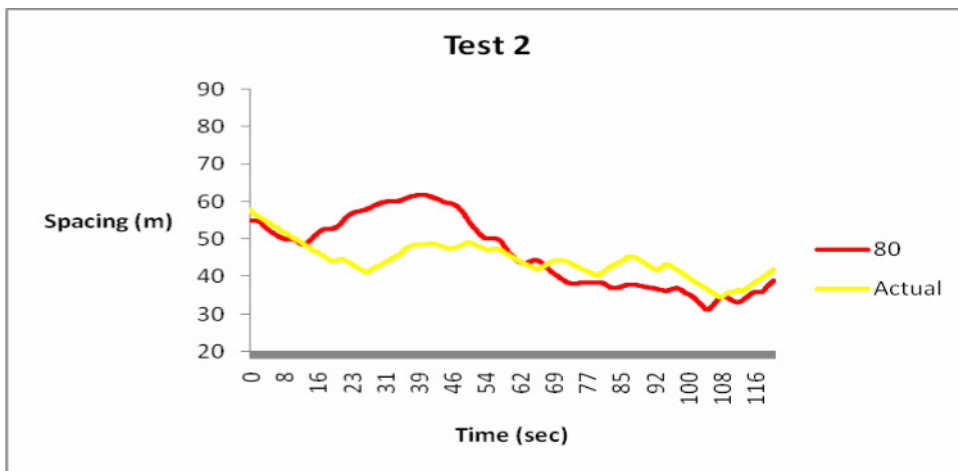
Figure 18 shows different threshold spacing values against actual space headway data. The 76 m and 80 m spacing thresholds are the closest to the actual curve. Figure 19 illustrates how the 80 m value is close to the actual data, while Figure 20 shows that 80 m is the optimum since it corresponds to the minimum value of RMSE.

To determine the optimum value for the two parts of this test (speed and space headway real data), the obtained results indicate that there are two optimum values: one from real speed data (86 m) and the second from real space headway data (80 m).

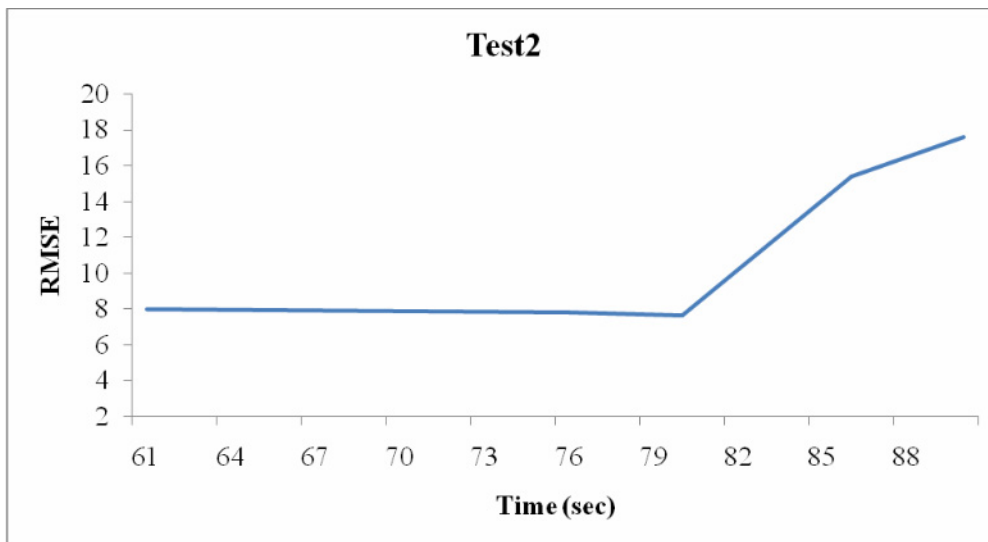
To discuss this difference, let 86 m spacing threshold be the optimum value. In fact, this value represents the minimum value in terms of real speed value, as shown in Figure 13. However, the difference between this and the minimum value of 80 m in terms of real space headway is 7.73 m, as shown in Figure 20. Therefore, the 80 m value represents the optimum value in terms of space headway real data. The difference between this value and the minimum value (optimum) in terms of real speed data is only 0.02 m/sec, as shown in Figure 13. This difference is trivial when compared with 7.73 m for the 86 m spacing threshold. Accordingly, the 80 m spacing threshold can be considered as the optimum value.



**Figure 18.** Different values of spacing threshold compared with the actual data.



**Figure 19.** Values of spacing threshold at 80 m compared with the actual data.



**Figure 20.** Different values of spacing threshold against RMSE values.

## 5. Conclusion

This paper identifies the limitations of the GHR model and focuses on the spacing threshold since it has a significant effect on the behaviour of the follower. The parameters of the GHR model ( $c$ ,  $m$ , and  $L$ ) have been calibrated and the best combinations are as follows:

For acceleration:	m = 1.0	L = 1.1	c = 1.2
For deceleration:	m = 1.1	L = 1.4	c = 1.2

The results from the two tests state that the 80 m spacing threshold is the most suitable value amongst the other values that were selected by different researchers. To get the best results for these parameters in the car-following model, more data is needed in the calibration process. Further data is needed to test different conditions, such as with stop-and-go conditions.

## References

- Aron, M.,(1988) Car-Following in an Urban Network: Simulation and experiments. In Proceedings of Seminar D, 16th PTRC Meeting (pp. 27-39).
- Aycin, M.F. (2001) *Development of a Car-Following Model to Simulate Driver and Autonomous Intelligent Cruise Controlled Vehicular Traffic Flow*. PhD Thesis, University of Illinois.
- Brackstone, M. and McDonald, M. (1999) 'Car-following: a historical review,' *Transportation Research* 2 (4), pp. 181–196.
- Chandler, R., Herman, R., and Montroll, E. (1958) 'Traffic dynamics: studies in car following,' *Operations Research* 6 (2), pp. 165–184.
- Edie, L. (1961) 'Car-following and steady-state theory for noncongested traffic,' *Operations Research* 9 (1), pp. 66–76.
- Gazis, D.C., Herman, R., and Rothery, W. (1961) 'Nonlinear follow-the-leader models of traffic flow.' *Operations Research* 9 (4), pp. 545–567.
- Gazis, D.C., Herman, R., and Potts, R. (1959) 'Car-following of steady-state traffic flow,' *Operations Research* 7 (4), pp. 499–502.
- Gipps, P.G. (1981) 'A behavioural car-following model for computer simulation,' *Transportation Research, Part B*, 15 (2), pp.105–111.
- Herman, R., Montroll, E., Potts, R., and Rothery, R. (1959) 'Traffic dynamics: analysis of stability in car following,' *Operations Research* 7 (1), pp. 86–106.
- Hoefs, D.H.,(1972) *Development of a method of Car-following* , University (TH) Karlsruhe, Germany, 1972.
- Ma, X. and Andreasson, I. (2006) 'Estimation of driver reaction time from car-following data: application in evaluation of general motor-type model,' *Transportation Research Record: Journal of the Transportation Research Board*. pp, 130–141.
- May, A.D., and Keller, H. E.,(1967) *Non-Integer Car-Following Models*.HRR 199.
- Olstam, J.J. and Tapani, A. (2004) *Comparison of Car-Following Models*. Swedish National Road and Transport Research Institute.
- Ozaki, H.,(1993) *Reaction and Anticipation in the Car Following Behaviour*. In *Proceedings of the 13th International Symposium on Traffic Flow Theory* 349-366.
- Pipes, L.A. (1967) 'Car-following models and the fundamental diagram of road traffic,' *Transportation Research* 1, pp. 21–29.
- Saechan, P. (2009) *Traffic Flow Dynamics Using Car-Following Models*. MSc Dissertation, University of Salford, unpublished work.

- Treiterer, J., and Myers, J.A., (1974) *The Hysteresis Phenomenon in Traffic Flow*. Transportation and Traffic Theory, New York.
- Sauer, C., Andersen, G.J., and Saidpour, A. (2004) *Detection and Avoidance of Collisions: The REACT Model*. Institute of Transportation Studies, University of California, Berkeley, USA.
- Toledo, T. (2003) *Integrated Driving Behavior Modeling*. PhD Thesis, Massachusetts Institute of Technology.

# Calibration of visual angle car following model based on real site traffic data

Jalal Al-Obaedi, *School of the Environment, University of Salford*

## Abstract

Traffic simulation models have increasingly been used due to low cost, time saving and ability to examine possible solutions for traffic-related problems without interrupting the real life traffic conditions during testing alternative scenarios. Recently, visual angle car following model has been studied and examined by the author for use in developing a micro-simulation car following model. The proposed model has been tested for different angular velocity thresholds against real published traffic data by developing a simple-one lane micro-simulation program. The results show that when the suggested angular velocity threshold of about 0.003 rad/sec is applied, the model will become able to replicate real life traffic movements. Root Mean Square and Error Metric statistical tests have been used to compare different selected angular velocity thresholds. A non parametric, Kolmogorov-Smirnov test indicates that the difference between the observed and simulated data is significant when the angular velocity threshold is below a value of 0.002 rad/sec. The aim of developing this car following model is to be used at a later stage in developing a micro-simulation model to represent traffic behaviour at motorway merges and to test the objectiveness of using ramp metering strategies on motorways.

## Keywords

Car following, visual angle, micro-simulation, car following calibration

## 1 Introduction

Traffic simulation models play a major role in allowing transportation engineers to evaluate complex traffic situations and recommending alternative scenarios. Such simulation models provide the opportunity to evaluate traffic control and design strategies without committing a lot of expensive resources (including time) which are necessary to implement alternative strategies in the field (Clark and Daigle, 1997). According to Kotsialos and Papageorgiou (2001) these models can be used for estimation, prediction and control related tasks for the traffic process. Moreover, computer simulation models can help in analysing every day's traffic management needs by looking at problems such as congestion and identify their sources.

The main components of any traffic simulation model are car following, lane changing and gap acceptance models. Car-following models describe the relationship between pairs of vehicles in a single lane. This relationship is represented by several mathematical models which basically describe the effect of the leading vehicle on its follower. The lane changing model represents the lateral movement for traffic movements. The feasibility of making a decision for lane changing is based on the availability of sufficient gap in a target lane. Usually, the availability of such gap is controlled by gap acceptance model.

Car following models are well described and classified in the literature (see for example, Brackstone and McDonald (1999) and Panwai and Dia (2005)). This paper will focus on calibration of the visual angle car following model proposed by Al-Obaedi and Yousif (2009a, b).

## 2 Visual angle car following models

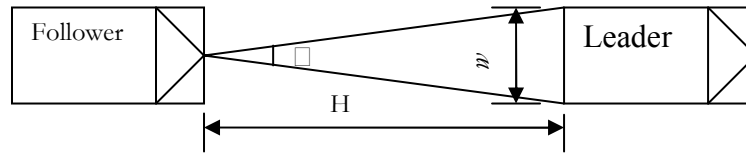
One of the earlier car-following models is the visual angle model. The visual angle as shown in Figure 1 is given by the following equation:

$$\theta = 2 \tan^{-1} \left( \frac{w}{2H} \right) \quad \dots \text{Equation 1}$$

Where:

$w$  is a width of the leading vehicle.

$H$  is the spacing between the leading and the following vehicles.



**Figure 1** Illustration of the visual angle ( $\theta$ )

Michaels (1963) observed that the detection of the relative velocity depends on the rate of change of angular motion (angular velocity) of an image across the retina of the eye of the follower driver (Fox and Lehman, 1967).

The angular velocity is found by differentiating Equation 2 with respect to the time ( $t$ )

$$\frac{\Delta\theta}{\Delta t} = -w \frac{V_L - V_F}{(X_L - \text{length}(l) - X_F)^2} \quad \dots \text{Equation 2}$$

Where:

$V_L$  and  $V_F$  are speeds of leading and following vehicles, respectively.

$w$  is the width of leading vehicle.

$X_L$  and  $X_F$  are positions of leading and following vehicles, respectively.

$\text{length}(l)$  is the length of the leading vehicle.

Visual angle models are described by previous researchers, such as Brackstone and McDonald (1999), and Panwai and Dia (2005), as one type of psychophysical or action point models since these models define the next vehicle's action on whether or not the follower exceeds certain thresholds. These assume fixed values (thresholds) for angular velocity. Once the absolute value of the angular velocity exceeds the threshold, the follower will accelerate or decelerate opposite to the sign of the relative angular velocity. Table 1 presents a brief summary of the values used for angular velocity thresholds by various researchers.

Recently, theoretical studies carried out by Al-Obaedi and Yousif (2009a, b) have argued that higher values for the angular velocity thresholds such as a value of 0.003 rad/sec as proposed by Hoffman and Mortimer (1994, 1996) are more reasonable. However, no real traffic data have been used in these studies. This paper tries to fill the gap in these studies through examining different angular velocity thresholds using published real site traffic data.

**Table 1** Summary of angular velocity thresholds used by various researchers

Researcher(s)	Threshold value ( $\Delta\theta/\Delta t$ ) (rad/sec)	Remarks
Michaels and Cozan (1963)	0.0003 – 0.001	Experimental
Fox and Lehman (1967)	0.0006	Simulation
Ferrari (1989)	0.0003	Simulation
Hoffman and Mortimer (1994 and 1996)	0.003	Experimental
Xin et al. (2008)	0.0008	Simulation

### 3 Visual angle model thresholds and assumptions

This section presents thresholds and assumptions for the visual angle model which are used in this study as described by Al-Obaedi and Yousif (2009a, b).

Four thresholds are used in the proposed model. Positive and negative angular velocity thresholds are used to arrange the difference in speed between the leader and the follower. While minimum and maximum time spacing thresholds (MinTH and MaxTH) are used to represent driver error in estimating his/her headway according to Weber's law (Brackstone, and McDonald, 1999)

The main assumption of the model is based on whether or not the angular velocity calculated from Equation 2 exceeds the assumed angular velocity threshold values. If the absolute angular velocity becomes higher than a certain selected threshold, the follower starts to accelerate or decelerate opposite in sign to that of the angular velocity value.

If the MaxTH threshold is exceeded, the follower will start applying acceleration to reach his/her desired headway. On the other side, if the MinTH is exceeded, the follower will apply deceleration in order to recover his/her desired headway.

If the angular velocity value calculated from Equation 2 is within the two visual angle threshold limits, and if the minimum and maximum time headway thresholds are not exceeded, the driver is assumed to keep a constant speed.

The selected values for acceleration or deceleration are the minimum of the following rates (see for example Fox and Lehman, 1967 and Ferrari, 1989):

- the acceleration rate which is required to reach the desired speed,
- the acceleration/deceleration rate required to reach the leader's speed, and
- the acceleration/deceleration rate required to maintain the desired spacing using the following Equation 3 below. This equation is derived based on the same assumptions reported by Hidas (1996).

$$ac(f) = \frac{[x(l, t + \Delta t) - x(f, t) - v(f, t) * \Delta t - DTHead(f) * v(f, t) - length(l) - Buf]}{\frac{\Delta t^2}{2} + DTHead(f) * \Delta t} \quad \dots \text{Equation 3}$$

Where:

$ac(f)$  is the acceleration (or deceleration) rate of the follower

$\Delta t$  is the scanning time.

$DTHead(f)$  is the desired time (spacing) for the follower.

$x(l, t + \Delta t)$  is the position of the leader at time  $t + \Delta t$

(other terms are as defined before).

$Buf$  is the required buffer spacing by the follower.

$v(f, t), x(f, t)$  are speed and position of the follower at time  $t$ , respectively.

#### 4 Calibration Methodology

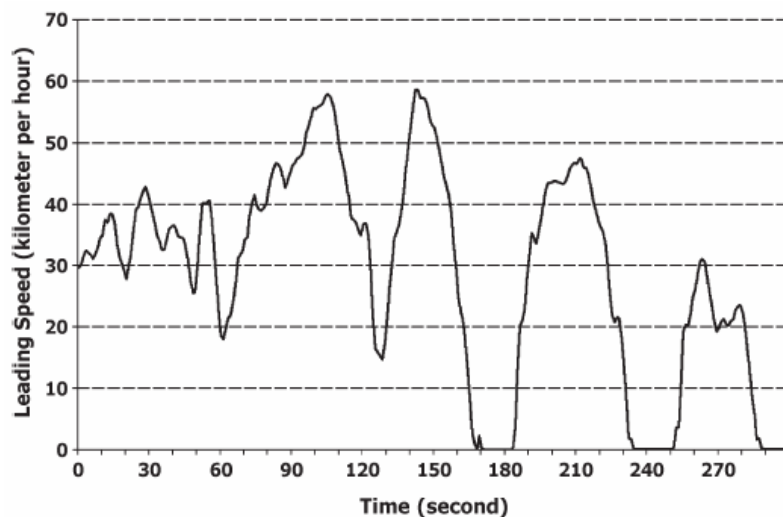
The reliability of any model depends on how well that model could represent the reality (Barceló and Casas, 2002). The calibration of simulation models is an iterative process to select the best parameters for a given model depends on real traffic data.

For this paper, real traffic data as reported by Panwai and Dia (2005) has been used to calibrate the visual angle model. The model parameters will be varied to find out the best fit for the data based on statistical tests. The following subsections explain the data, statistical

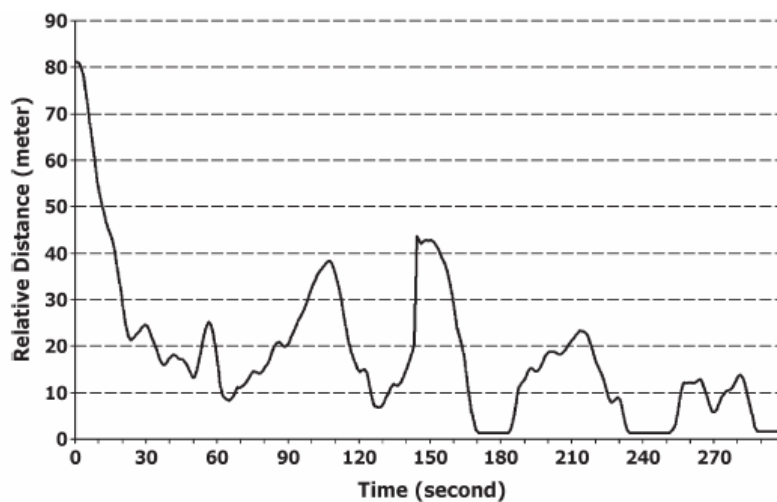
tests that used, and the calibration parameters. A micro-simulation program has been prepared as a bed test for this study.

#### 4.1 Data description

The source of the data used in this paper is taken from Panwai and Dia (2005) which is based on two vehicles trajectories while these vehicles are travelling at stop-and-go conditions for a distance of 2.5 km for 300 seconds. Figure 2 shows the speed profile for the leading vehicle while Figures 3 and 4 represent the clear spacing and relative speed profile between these two vehicles respectively. The speed range was between 0 and 60 km/hr. As shown in Figure 2, both vehicles came to full stop several times during the whole period. For the purpose of this research, numerical values for the leading speed from Figure 2 and the clear spacing from figure 3 are abstracted for each 0.5 seconds interval.

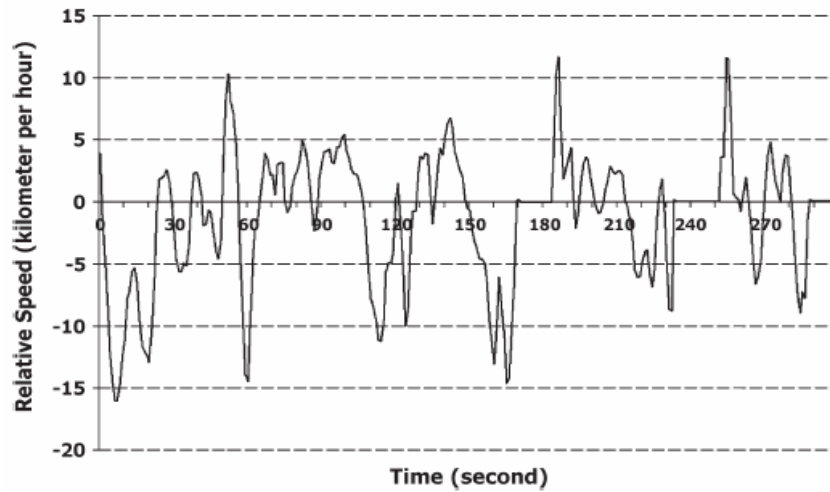


**Figure 2** Leading vehicle profile (source: Panwai and Dia, 2005)



**Figure 3** Clear spacing between the two vehicles (source: Panwai and Dia, 2005)





**Figure 4** Relative speed profiles between the two vehicles (source: Panwai and Dia, 2005)

According to Panwai and Dia (2005) this data has been used to evaluate the behaviour of different micro-simulation models. Table 2 shows a summary of some of the work done using this.

**Table 2** Summary of car following models which were evaluated using the data

Model	Statistical tests		Reference
	Root mean square error	Error metric	
MITSIM	Not used	3.75	Manstetten et al. (1997)
Wied/Pel	Not used	14.01	Manstetten et al. (1997)
Wied/VIS	Not used	10.67	Manstetten et al. (1997)
NSM	Not used	24.51	Manstetten et al. (1997)
OVM	Not used	9.37	Manstetten et al. (1997)
T <sup>3</sup> M	Not used	2.4	Manstetten et al. (1997)
AIMSUN (v4.15)	4.99	2.55	Panwai and Dia (2005)
VISSIM (v3.70)	5.05	4.78	Panwai and Dia (2005)
PARAMICS (v4.1)	10.43	4.68	Panwai and Dia (2005)

## 4.2 Statistical tests

Root mean square error (RMSE) and Error metric (EM) statistical tests (see Equations 4&5) have been used in this study to estimate the error value for the actual and simulated clear spacing between the two vehicles. Moreover, Kolmogorov-Smirnov non parametric test has been used to make a decision on whether to accept or reject the hypothesis that observed and simulated data (in term of clear spacing) are significantly different.

$$RMSE = \sqrt{\frac{(ds - df)^2}{n}} \quad \dots\text{Equation 4}$$

$$EM = \sqrt{\left(\log \frac{ds}{df}\right)^2} \quad \dots\text{Equation 5}$$

Where:

$ds$  is the simulated spacing between two vehicles (m)

$df$  is the actual spacing between two vehicles (m)

### 4.3 Model parameters

Some of the model parameters could be directly estimated from the data. These include the desired speed and the buffer spacing for the follower. The desired spacing of 60 km/hr has been assigned for the follower representing the maximum follower's speed during the period of 300 seconds. For the buffer spacing, a value of 1.5 m has been chosen representing the minimum spacing between the two vehicles at stopping conditions.

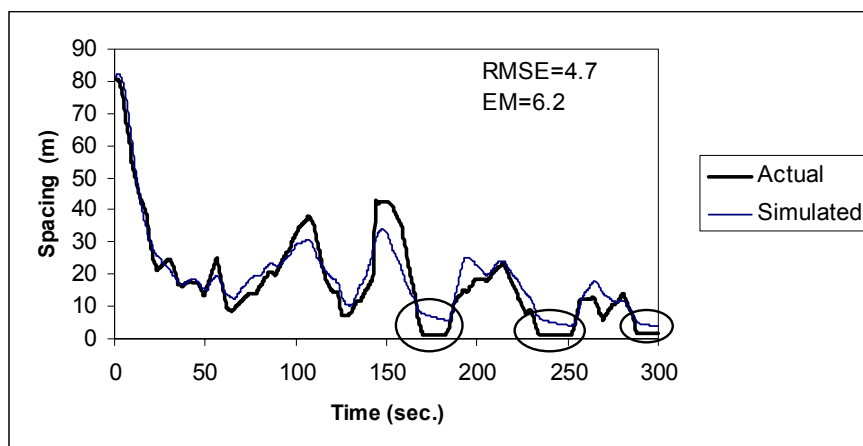
The desired time headway  $DTHead$  given in Equation 3 is chosen as 1.6 sec. based on several iterations to select this parameter. The selected parameters for calibration are the absolute angular velocity threshold  $\theta$ , minimum (MinTH) and maximum (MaxTH) time spacing. Angular velocity thresholds values from 0.0001 to 0.006 rad/sec are used. Values of  $(1.12*DTHead, 1.2*DTHead)$  and  $(0.88*DTHead, 0.8*DTHead)$  are selected for each minimum and maximum time headway thresholds, respectively to represent the effect of just noticeable difference according to Weber law (Brackstone and McDonald, 1999). Table 3 represents the combination of these parameters used in the calibration process.

**Table 3** Combinations of the calibration parameters

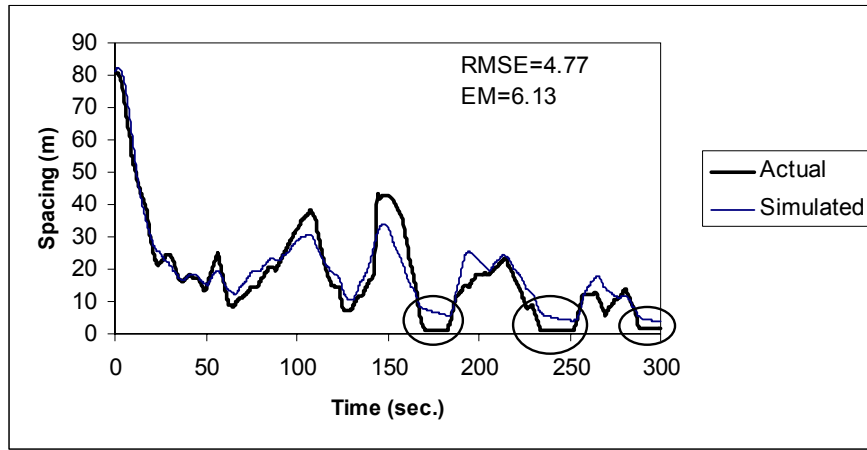
Case	Selected $\theta$ values (rad/sec)	MinTH (sec)	MaxTH (sec)
A	0.0001, 0.0003, 0.0008, 0.001,	0.88*DTHead	1.12*DTHead
B	0.002, 0.003, 0.004, 0.005 and 0.006	0.8*DTHead	1.2*DTHead

## 5 Results and discussion

Figures 5 and 6 represent the actual and simulation spacing between the two vehicles for Cases A and B with angular velocity threshold of 0.003 rad/sec. Although both figures show good agreement between simulated and observed spacings depending on RMSE values, the EM values as shown are too high compared with values in Table 2. Not like RMSE, the EM depends on the ratio of simulated to observed values as shown in Equation 5 and therefore, the higher values of EM for the cases in Figures 5 and 6 are due to stopping conditions (see circled parts in Figures 5 and 6). Therefore, further modification is required to the visual angle model assumptions relating to stopping conditions.

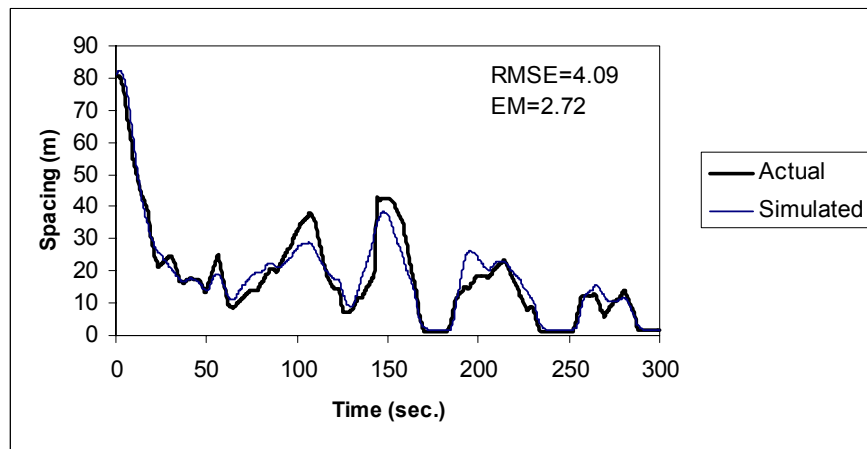


**Figure 5** Actual and simulated spacing for case A with  $\theta$  of 0.003 rad/sec

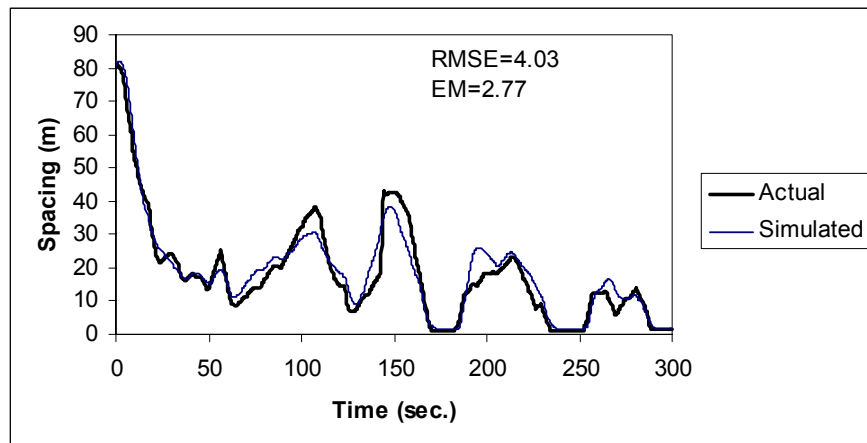


**Figure 6** Actual and simulated spacing for case B with  $\theta$  of 0.003 rad/sec

The suggested modification states that at slow speeds (up to 25 km/hr based on sensitivity analyses for this factor), drivers will not tend to react to the difference in speed and the main goal at this range of speed is to keep a minimum buffer distance. In the model and when the speed is less than 25 km/hr, the acceleration of the follower is assumed to be the minimum acceleration to maintain the desired speed or to reach the desired headway from Equation 3. The effect of this new assumption on the results for the same angular velocity threshold value of 0.003 rad/sec is shown in Figures 7&8 where both RMSE and EM seem to be within acceptable values. Moreover, the RMSE is found to be less than those in Table 2 which indicate that the visual angle model is able to replicate real traffic movements and therefore, the rest of work in this paper will be based on this assumption.

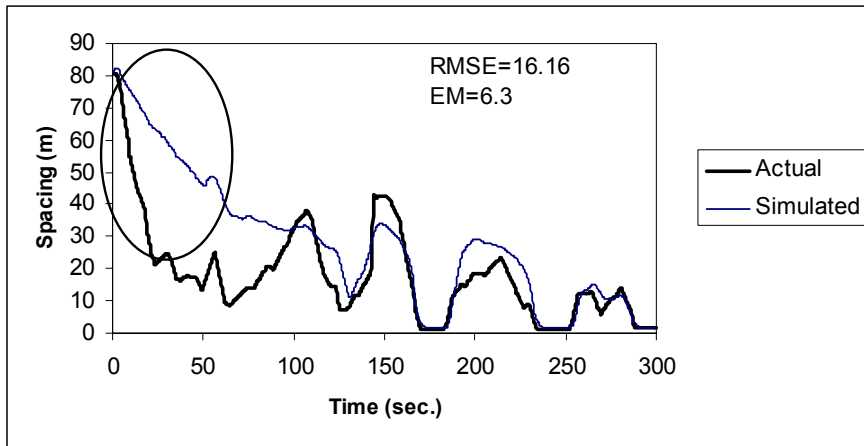


**Figure 7** Actual and simulated spacing for case A with  $\theta$  of 0.003 rad/sec

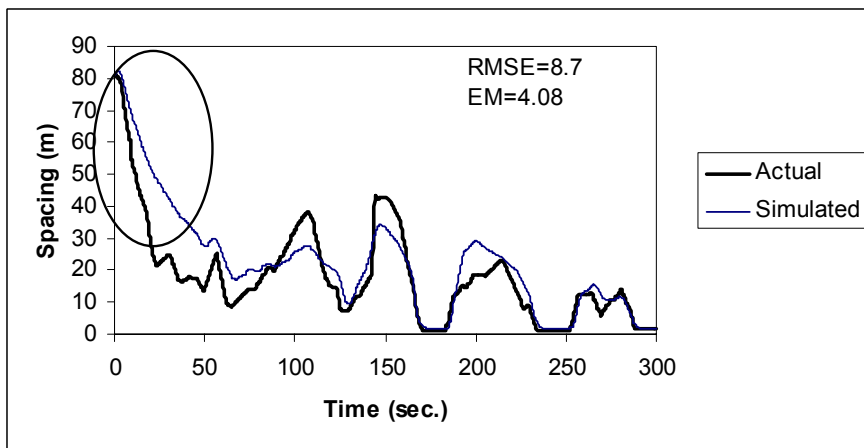


**Figure 8** Actual and simulated spacing for case B with  $\theta$  of 0.003 rad/sec

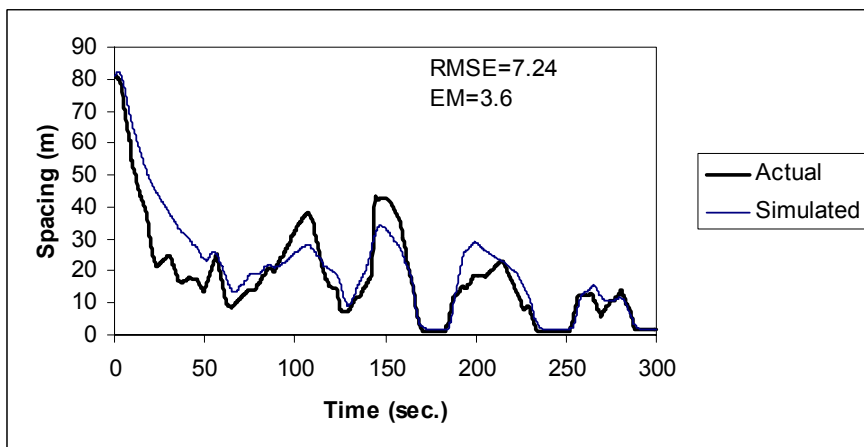
Also, there is no significant difference between Cases A and B. Therefore, it is decided to test other values of the angular velocity threshold depending on case A only. Figures 9-15 represent the actual and simulated spacing for different angular velocity thresholds for Case A only.



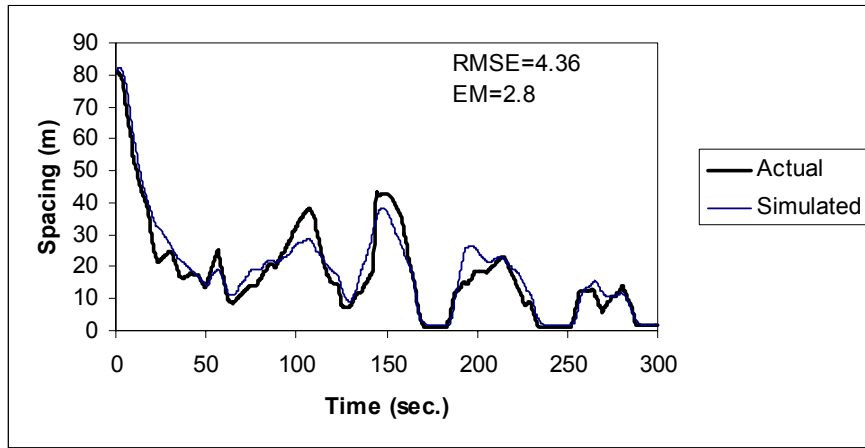
**Figure 9** Actual and simulated spacing for case B with  $\theta$  of 0.0003 rad/sec



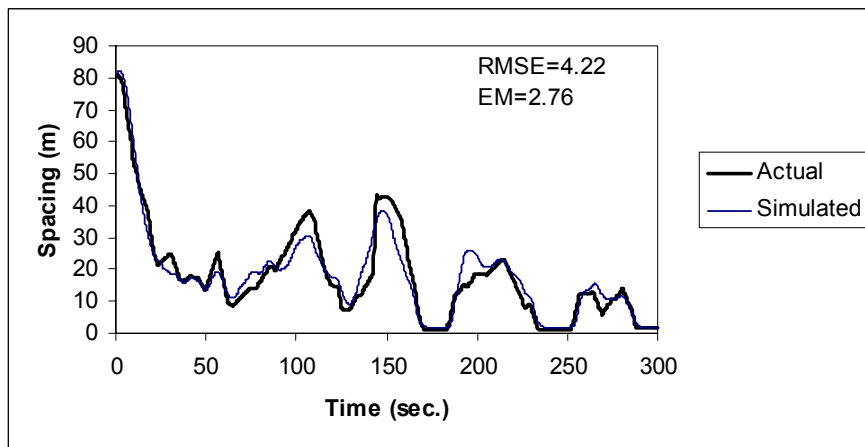
**Figure 10** Actual and simulated spacing for case B with  $\theta$  of 0.0008 rad/sec



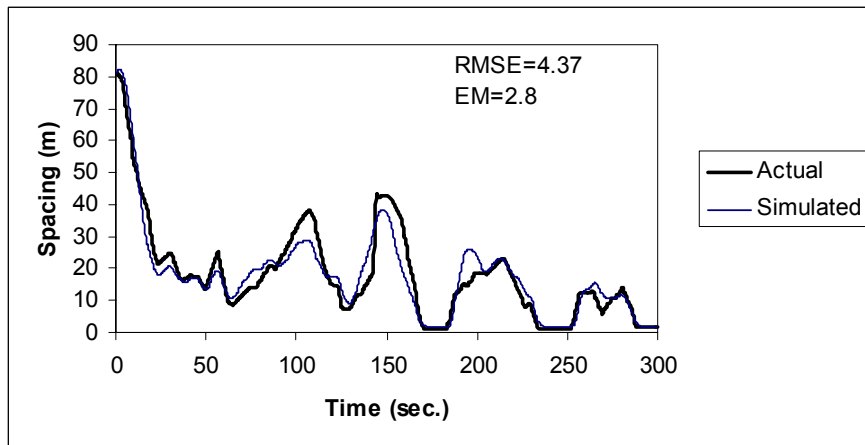
**Figure 11** Actual and simulated spacing for case B with  $\theta$  of 0.001 rad/sec



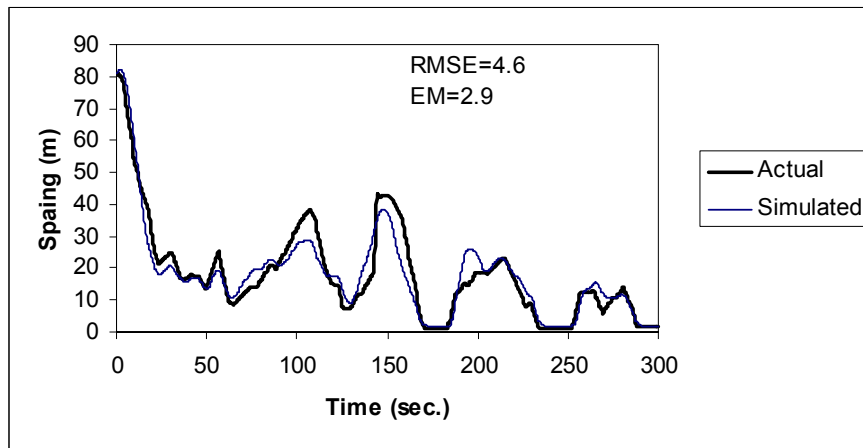
**Figure 12** Actual and simulated spacing for case B with  $\theta$  of 0.002 rad/sec



**Figure 13** Actual and simulated spacing for case B with  $\theta$  of 0.004 rad/sec



**Figure 14** Actual and simulated spacing for case B with  $\theta$  of 0.005 rad/sec



**Figure 15** Actual and simulated spacing for case B with  $\Theta$  of 0.006 rad/sec

The figures show that only when the angular velocity is 0.002 rad/sec or higher, there will be good agreement between the real data and the model. As shown in Figures 7 and 8, the minimum values for RMSE are 4.09 for Case A and 4.03 for Case B with angular velocity threshold of 0.003 rad/sec. Compared with other car following models, these values (i.e. 4.03 and 4.09) are found to be lower than all values reported in Table 2.

The problem in applying lower values of angular velocity thresholds such as a value of 0.0003 rad/sec as used by Fox and Lehman (1967) and Ferrari (1989) or a value of 0.0008 rad/sec as used by Xin et al. (2008) is that the follower will apply deceleration rate even when the distance between the two vehicles is too high. This behaviour is shown in the circled part of Figures 9 and 10 for the angular velocity threshold of 0.0003 rad/sec and 0.0008 rad/sec respectively.

## 6 Hypothesis testing

In order to find whether or not the difference between the simulated and real spacings is significant, Kolmogorov-Smirnov test has been used. According to this non parametric test, a hypothesis is accepted (i.e. the difference is insignificant) if the maximum difference ( $D_{max}$ ) in the cumulative probability is less than the critical limit ( $D_{cr}$ ). If the difference is higher than that limit, a hypothesis is rejected (i.e. the difference is significant). Table 4 represents a summary of applying this test for the 95% confidence level.

**Table 4** Hypothesis testing summary

Angular velocity threshold	RMSE	EM	G-S test		Hypotheses
			Dcr	Dmax	
0.0003	16.16	6.3	0.055	0.133	Rejected
0.0008	8.7	4.08	0.055	0.085	Rejected
0.001	7.24	3.60	0.055	0.063	Rejected
0.002	4.36	2.80	0.055	0.028	Accepted
0.003	4.09	2.72	0.055	0.034	Accepted
0.004	4.22	2.76	0.055	0.035	Accepted
0.005	4.37	2.80	0.055	0.036	Accepted
0.006	4.6	2.9	0.055	0.037	Accepted

The Table shows that the hypothesis is only accepted when the angular velocity threshold is about 0.002 rad/sec or higher. This confirms the theoretical work by Al-Obaedi & Yousif (2009a, b) in suggesting the use of higher values for the angular velocity thresholds in visual angle car following models.

## 7 Conclusion and Further Research

Visual angle car following model has been calibrated using real site data based on two vehicles trajectories as reported by Panwai and Dia (2005). The main finding of this study is confirming previous theoretical work by Al-Obaedi & Yousif (2009a, b) in suggesting the use of higher values for the angular velocity thresholds in car following models than used in the past in simulation applications.

It was found that using of angular velocity threshold of 0.002 rad/sec. and higher gives good replication of the real data. Moreover, the RMSE from the visual angle model is less than that reported in previous research work for other car following models using the same data.

Further research is needed to examine the ability of the visual angle car following model to replicate real traffic movements according to types of vehicles (i.e. Car Following Car, Car following HGV, HGV following HGV and HGV following Car).

## References

- Al-Obaedi, J. and Yousif, S. (2009a) Examining the limitations for visual angle-car following models, UTSG 41<sup>st</sup> annual conference, UCL.
- Al-Obaedi, J. and Yousif, S. (2009b) The use of visual angle in car following traffic micro-simulation models, International Built & Human Environment Research Week (IRW).
- Barceló, J. and Casas, J. (2002) Dynamic network simulation with AIMSUN, "Dynamic network simulation with AIMSUN," in Int. Symp. Transport Simulation, Yokohama, Japan, 2002, pp. 1-25.
- Brackstone, M. and McDonald, M. (1999) Car following: a historical review. *Transportation Research Part F*, Vol 2, pp. 181-196.
- Clark, J. and Daigle, G. (1997) The Importance of simulation Techniques in ITS research and analysis. *Proceedings of the 1997 Winter Simulation Conference. USA*.
- Ferrari, P. (1989) The effect of driver behaviour on motorway reliability. *Transportation Research Part B*, Vol 23B, No. 2, pp. 139-150.
- Fox, P. and Lehman, F.G. (1967) *Safety in car following*. Newark, (N.J.), Newark college of Engineering.
- Hidas, P. (1996) A car-following model for urban traffic simulation. *Traffic Engineering and Control*. Vol 39, pp. 300-305.
- Hoffman, E.R. and Mortimer, G.R. (1994) Estimate of time to collision. *Accident Analysis and Prevention*, Vol 26, Issue 4, pp. 511-520.
- Hoffman, E.R. and Mortimer, G.R. (1996) Scaling of relative velocity between vehicles. *Accident Analysis and Prevention*, Vol 28, Issue 4, July 1996, pp. 415-421.
- Kotsialos, A. and Papageorgiou, M. (2001) The Importance of traffic flow modelling for motorway traffic control, *Networks and Spatial Economics*, 1, Netherlands, pp. 179-203.
- Manstetten, D., Krautter, W. and T. Schwab, "Traffic simulation supporting urban control system development," in 4th World Congr. Intelligent Transport System, Berlin, Germany, 1997, pp. 1-8.
- Michaels, R.M. (1963) Perceptual Factors in Car following. *In Proceeding of the 2nd International Symposium on the Theory of Traffic Flow - OECD*. pp. 44-59, Paris.
- Michaels, R.M. and Cozan, L.W. (1963) Perceptual and field factors causing lateral displacements. *Public Roads* 32. pp. 233-240.

- Panwai, S. and Dia, H. (2005) Comparative evaluation of microscopic car-following behaviour. *IEEE Transactions on Intelligent Transportation System*, Vol 6, No. 3, pp. 314-325.
- Parker, M.T. (1996) The effect of heavy goods vehicles and following behaviour on capacity at motorway roadwork sites. *Traffic Engineering and Control*, Vol 37, pp. 524-531.
- Xin, W. Hourdos, J., Michalopoulos, P. and Davis, G. (2008) The less-than-perfect driver: A model of collision-inclusive car following behaviour. Transportation Research Board Annual Meeting.



# Changing Understanding of Emotion in an Educational Context

**David Menendez Alvarez-Hevia**, *Education and Social Research Institute, Manchester Metropolitan University*

## Abstract

The view of emotional phenomena within the educational context has changed significantly during recent decades. Different approaches to social sciences have provided new ways to interpret the emotional interactions between educational agents, their context and other aspects embedded in education. From a socio-critical approach this theoretical paper provides a review of different theories and ideas to bear in mind when it comes to researching aspects of emotions in an educational contexts. Hence, issues such as "emotional intelligence" (Goleman, 1996), "emotional labour" (A. Hochschild, 2003) and the social construction of emotions (R.Harrè, 1986) are explained and discussed throughout this work.

**Keywords:** Educator's emotions, emotional intelligence, emotional labour, *special* schools.

## Introduction

Mainstream schools and their teachers, specialised in subjects, are the dominant body of educational agents; nevertheless they are not the only habitants of the institutionalised educational space. Coexisting with them are other professionals working in centres which are considered "special" schools, for the reason that the nature of their students is far from the idealised norms dictated by developmental theories inscribed with the national curriculum. At first glance, the interaction between these two types of school is presented in education as a complementary alternative in equilibrium. The relationship, however, is not as equilibrated and transparent as it might appear.

Postructuralist theory challenges the stability of binaries, opening and transforming spaces within the discourses of ideas (Jencks, 1995; Drolet, 2004; Check and Gough, 2005). Rather than see this coexistence between mainstream and special schools as forged in equal terms, it can be seen as a binary opposition. According to Derrida, the gain of the presence of one reality is conditioned by the existence of another interlinked reality which is disqualified by the dominant one (Derrida, 1976). Without this type of subjugation, the space necessary for the presence of mainstream schools cannot be opened and therefore the existence of these "conventional" schools is challenged. As pointed out by many authors, this binary opposition is an indispensable condition to bear in mind when carrying out discourse analysis (Gee, 1999). Critical theories such as feminism, which focuses its approach on the deconstruction of the gender dualism female/male (Plumwood 1993; Smith 1999; Hughes 2002), or race theory, centring on issues related to racial differences (Ladson-Billing, 1995; Parker and Roberts, 2005), are some of the most common examples of the potential of questioning the dualism for qualitative inquiry, within the appearance of deconstruction or discourse analysis.

In the field of education, the presence of this hostility is also confronted. MacLure (2003) illustrates some of these oppositional poles in, what is ironically called by Gloria Ladson-Billing (1995), "the nice field of education", with reference to the illusion of equality and a peaceful sense that apparently reigns in this domain. This misconception conceals a completely different image where the will to power drives movements to advantaged positions. Thus, as an example, we can find oppositions between discourses about the inside (classroom)/outside (world), as opposed spaces where knowledge is constructed, acquired and validated in different modes (p.13). Another example of the existence of this dualism in education comes from the polemic use of standard English/non-standard English in the school as a way to privilege the language spoken by educated middle class speakers (p.22).

This research aims to take into account the dualistic relationship between mainstream schools and "special" schools without comparing them. Previously, studies have generated relationships among different categories with the purpose of allocating differences between the two positions. This research is based on the examination of two different cases of education

professionals working in these so-called "special schools", focusing on their emotional interactions; in order to stimulate an open debate about the diverse emotional implications among different types of educational centres.

### **Researching emotions: an opportunity for a real change in education, or just an illusion?**

Emotion provides a new perspective for research in social sciences in these changing times (Stearns & Stearns, 1988). However, education is not changing at the same pace as our society; it is developing many strategies of resistance to change, which results in an old fashioned culture and practice, as Hargreaves (1998) points out. Some of these components which sustain this model of resistance are hidden in the label "new orthodoxy", for example; centralised curriculums, indicators and rubrics of achievement, deeper learning based on memorisation or rigid learning standards for everybody (Hargreaves, Moore and Manning; 2001). They do not, however, achieve the introduction of a real change; they only serve to maintain the same model of school, education and consequently of society... a "wolf in sheep's clothing".

New issues introduced in education, such as the interest in aspects related to the emotional world of students in the first place and educators in the second place, although offering some benefits, are also suspected of being revised by researchers who can see these as delicate innovations (see for example Boler, 1999). This new interest in emotions gives a new sense to the social world enlightened by reason. Attention to emotions provides new spaces and possibilities for the study of the relationships and interactions of professionals working in education. It is essential to take a look at their life stories and their construction of themselves as educators to understand their circumstances as a whole.

According to Clarke, the "situation is always greater than the sum of its parts because it includes their relationality in a particular, temporal and spatial moment" (Clarke, 2005: 23). For that reason I propose to take a look at the emotions of these types of professionals working in these "special" circumstances, to draw a more complete picture of them. However, to carry out this task it is necessary to take into account their position to challenge the assumption that their experience is the same as that of education professionals working in mainstream schools. Taking into consideration other peoples' standpoint is an idea critiqued by researchers from many disciplines within social sciences, see for example Fine and Weis (1998); Atkinson, Coffey et al (2001) and Coffey (2002). This open debate helps to expand the forms of social research, blurring and expanding the boundaries of the methodology to the point that as in this research, ethnography and case study take the form of a combined open approach.

### **Narratives from the field: exploring emotional involvement**

The narratives from the field serve to draw a picture of the emotional involvement of the people from the centres concerned with the research. Nevertheless, I agree with the idea that this re-telling I have done as a researcher is problematic, with ethical consequences, even when I am talking about my own experience since "the norms by which I seek to make myself recognisable are not fully mine", they are in a continuous conflict between the self and the social and historical temporality (Butler 2005: 35).

I am interested in narratives from the field which are in line with the ideas of Jean-François Lyotard's (1984) essay "The Postmodern Condition", which proclaimed the disappearance of the great master narratives after their failure in achieving the modern objective of creating fixed and stable general, social and natural truths. Instead of total disappearance, he argues that they are buried in an "unconscious existence" highlighting more local narratives which serve to interpret the world in a different way. Local and everyday discourse is part of the postmodern turn which casts doubt on what is considered knowledge, locating an inseparable relationship with power (Foucault, 1991). Therefore, the stories captured in the field alone are important, generalisation is not the aim because it is problematic (Roth, 1987); these narratives are not limited to being analysed or theorised, and actually they can be re-told and long used (Ellis & Bochner, 1996). The emotional experience narrative of education professionals is dynamic, it is immersed in a continuous movement which is caused by

different interactions. Because of this instability, narratives must be open to new interpretations, where Friedrich Nietzsche and later Michael Foucault proclaimed the "death of the author" as a mode to liberate the text from the limitations that are constraining the potential of opening it and its possibilities (Foucault, 2000).

The access to the narratives is obtained in different ways to combine the written and spoken word, making an attempt to break the tendency of situating the second closer to the truth or logos of meaning. Both, combined, are important for the development of language (Derrida, 1976). Even if language gives us only traces of reality, and is moreover not innocent, and, additionally, does not give access to an objectively graspable reality (MacLure, 2003), it does not mean that we, as researchers, are completely unarmed to face reality. Playing with the metaphor of the Greek *Pharmakos* used by Derrida to deconstruct the differences between speech and writing (Mortensen, 2000), I support the idea that language is what we have for research, a double-edged sword that cures and kills, saves and condemns, cleans and tarnishes, all at the same time, again and again, over and over. On the one hand, it is our cure in the form of creator of an illusion of access to reality to satisfy our "*desire of presence*" (Derrida, 1997: 143). On the other hand, it is a poison that pollutes our access to reality creating more disquiet and gaps between what is inside and outside. Nevertheless, we should bear in mind that far away from language - we are astray and vanished into a chaotic world without meanings. It is also expressed by Derrida in his renowned quotation: "There is nothing outside of the text [there is no outside-text; il n'y a pas de hors texte]" (1997: 158). In other words, language constructs what we consider reality, shaping and giving sense to our discernment about the world. Therefore escape from language, or only focus in language as an innocent cognitive process, or an arbitrary creation subsequent to thought, or simply a tool to communicate between the world inside the subject (self) and the world outside the subject (social), supposes a risk for the postmodernist researcher, who takes a new standpoint to consider language and discourse.

Talk about text or language within a poststructuralist approach, is talk about discourse. A definition of what it is can be problematic because it takes different forms according to different authors (Wetherell, Taylor and Yates, 2001). Therefore, what we can find through these different approaches to the discourse, are diverse forms of discourse analysis as a qualitative methodological tool at the disposition of the researcher of social sciences. The significance of the discourse for the post structuralist theorist is defined by Pathy Lather as so extreme that it "worlds the world" (Lather 1993: 675), hence it drives our ways of knowing and proceeds in a continuous battle for reigning spaces. It is very important to comprehend the discourses that shape special needs education (Allan, 1999) and, furthermore, other "special" education the emotional experience narratives of educational professionals.

### **The emotional discourse**

In particular, in the case of this research, the education professionals' emotional discourse can provide information about how the emotional subject is (or is not) established, how they deal with their emotions in different aspects of their life and what agents and discourses are present. Different approaches to discourse are taken into account throughout this research, especially the contribution of Foucault in what is known as his "box of tools" (Pastor & Ovejero 2006: 31). Through his work he combines archaeology and genealogy to develop a form of discourse analysis focused on the formation of discursive practices, power relations and relations of selves. In this Foucauldian line of argument, teachers' emotions can be conceptualised as discursive practice: "they are not private reactive response to events but are socially organised and managed through social conventions, community scrutiny, legal norms, familiar obligations and religious injunctions" (Rose 1990: 1 in Zembylas, 2005: 936). Moreover, to work with the emotional discourses of the education professionals, focusing on how specific stories are constructed, requires a vital analysis of the links between emotions and situations (Edwards, 1999), giving special attention to the relationships between the different subjects involved. For a better understanding of this reality, Adele Clarke (2005) recommends the analysis of the differences through three different maps; situational, social worlds and positional maps. The author combines different discourse analysis trends, to develop a very practical and understandable innovative form of grounded theory in the line of postmodern ideas. It is very significant that one of the cases analysed in her book talks about a study of emotions - in this case, instead of the field of education, it is in the field of health nurses (see Bone, 2002).

## **An open debate about emotions: a field of dispute**

This debate concerning emotions and discourse calls for more attention to the field of education, within an environment previously defined as detached from the mainstream idea of education catalogued as "special". In addition, it is necessary to define the ideas, theories or body of knowledge which shapes the vision of the concern of this work: the emotions of the education professionals. Focusing on recent studies about emotions, its concept and notions are examined alongside many disciplines such as philosophy (Goldie, 2000), psychology (Griffiths, 1997) and sociology (Scheff, 1990). Conversely, I would like to emphasise the idea that the current study of emotions itself entails an overlapping of disciplines. For example "modern science" is highly criticised since it acts as a repressor of new possibilities, imposing its own view of the world in a way that tyrannise other forms and modes ( Foucault, 1989) - in other words, what it does is assemble facts through the practice of "splitting and inversion"(Woolgar, 1988). Some of the radical, negative consequences of this are considered to be "alienation, dehumanisation, ecological degradation and nuclear, chemical and biological hazards and warefare" (Croysant & Restivo 1995: 70) and, also mentioned, are the subtle, but active, suppression, marginalisation and restraining of access of some subjects according to their gender, race or other categories. This repression takes the form of assumption, such as the idea of women as a subject more likely to be a slave to their emotions; furthermore, different medical and biological arguments are exposed to support this idea. The wicked intention of this close connection between emotion and women is none other than that of maintaining the primacy of men in the science through the idea that scientific knowledge is free of emotional contamination. Hence, an emotional subject is undesirable (Boler, 1999).

This limitation of access is constructed through relations of power by the scientific assumption of objectivity, impartiality and realism, as well as constituting the basis to maintain its primary discourse positioning. Power should not be seen only as a negative instance because it is also a productive network running around the whole social body (Focault 1995b: 61). The work of Paul Feyerabend (1975) aims to deconstruct the scientific framework through "theoretical anarchism" that is, a proposition that entails a fierce opposition to the universal standards that restrict creativity and free change. Feyerabend believes in an alternative approach "more humanitarian and more likely to encourage progress" (p.5)

## **Rethinking the scientific position of emotions**

In the midst of this split in the stability and status of scientific recognition, new spaces to research emotions in the social sciences are open, whereby disciplines have to be restructured. We have scientific disciplines like neuroscience or cognitive psychology that, although maintaining their approach within the traditional physical and life sciences, in their later trends commence to call attention to new explanations out of this internal space determined by the experimental neuropsychology theories. One of the most significant authors in this field is Le Doux (1999) who situates the emotional activity in a part of the brain called the "tonsil" (a part of the limbic system) and rational activity in the neocortex. Both maintain a relationship when they come to process stimuli and give a response, making the emotional process faster but less elaborated than the rational process which gives a more socially accepted reaction. It brings to mind again the proposed question asked by James (1984): "whether feelings cause emotions or physical responses cause feelings, in other words: do we feel sorry because we cry? Or, do we cry because we feel sorry?" Although the question was explained by him according to the first statement which argues that emotional feelings follow bodily arousal, Cannon (1927) leaves the debate open again.

In addition, the neurophysiologist Damasio's book *Descartes' Error* (1995) explains the indivisible relationship between the mind situated inside (biochemical reactions, neural circuit, brain division etc.) and the body allocated outside (environment, social constructions etc.), seeing both as composing an indissoluble organism. He gives evidence of this tricky division between mind and body to justify a reformulation of one of the main assumptions behind the philosophical pillars of the scientific approach, summarised in the Latin quotation ascribed to the rationalist philosopher Descartes: *cogito ergo sum* – I think, therefore I am. Damasio challenges this assertion, first with the title of his book as a magnet for the critical reader with his attack on the philosopher, followed by the ideas contained inside. He proposes a shift in the direction of Descartes' idea, instead of seeing thought as a cause of existence; he sees

existence as a cause of thought, giving importance to the emotional world in a new way, where biochemistry and the neurological theories do not have to have the last word.

### **The emotional intelligence proposal**

The discourse surrounding emotional intelligence is of great importance for education since it is based on the possibility of development of the emotional skills, giving special importance to the social variables and education (Gardner, 1989; Goleman, 1996). The idea of emotional intelligence is constructed around five competences, abilities or skills based in the works of Mayer, Salovey and Gardner (Salovey and Mayers, 1990; Gardner, 1995). These four abilities that compose the corner stone of this theory are summed up below:

1. Self-awareness — the ability to read one's emotions and recognise their impact while using gut feelings to guide decisions.
2. Self-management — involves controlling one's emotions and impulses and adapting to changing circumstances.
3. Social awareness — the ability to sense, understand, and react to others' emotions while comprehending social networks.
4. Relationship management — the ability to inspire, influence, and develop others while managing conflict.

(Bradberry and Greaves, 2005)

However, this discourse, based on a cognitive approach, has its detractors. From critical perspectives such as feminism, emotions are seen as a "primary site of social control; a site of political resistance ...[which]... can mobilise social movements of liberation" (Boler 1999:13) and the emotional intelligence models are seen as a "reflection of a contemporary example of 'pastoral power'" whereby the "individual is seduced to police his/her emotions in the interest of neoliberalism, globalised capitalism" (Boler 1999: 22). For that reason, the model of emotional intelligence is seen as a mechanism of control that confines and limits the construction of subjects.

### **Social considerations**

The work of Hochschild (2003) concerns the risks that involve the social depreciation of the role of emotions within the development of some professions. It is a key reference work when discussing the social influence of the emotions and can be taken into consideration when we research the emotional involvement of special needs education professionals. The conditions described by Hochschild for jobs that can be considered as "emotional" are listed below:

1. *Involve face-to-face or voice-to-voice contact with public* (teachers actually do this since their everyday work is in direct interaction with students, parents, colleagues etc.).
2. *Require the worker to produce an emotional state in the other person.*
3. *Allow the employer, through training and supervision, to exercise a degree of control over the emotional activities of employees.*

What these types of explanation are making evident is the importance and necessity of understanding the social world if we want to develop and get an understanding of what emotions are. Nevertheless, from a relativist point of view, what we are going to develop is not more than a representation of what emotions are - in essence, traces of something to which we cannot get access. It is something that is not given to us as final universal truth but which we are still discovering. Maybe it is time to seriously consider whether reality is not as easily accessible as previously thought. However these traces are there, they can be analysed not as something waiting to be found, but rather, as representations in the form of constructions or discourses that drive our understanding. Hence, with regard to the emotional issues, the "rhetoric of the realism" that aspires to know the reality beyond constructive descriptions (Edwards, Ashmore and Potter, 1995) cannot give us more than another representation of this outside the social constructions. It is an incomplete picture, an impossible task made possible through the illusion of pure knowledge.

To talk about emotions in a different way that allows us to draw an improved picture, we should see them as a social construction subject to negotiation; in other words, as something not naturalistic because it is a concept in everyday use. It means that emotional concepts and

constructs should be considered within the standpoint of the language and the situation (Crespo, 1986), constituting a structured narrative (Goldie, 2000). Social constructionism through discourse theory is against the scientific interest in the universalisation of physiological reactions of emotions and its interest in revealing the existence, nature or working of hidden cognitive mechanism. In place of this unattainable exploration, Social constructionism proposes a focus on the discursive structures that give birth to the emotional world (Harre and Stearn, 1995).

## Conclusion

The analysis of the emotional culture of education professionals focuses attention on the social patterns which strengthen or promote the manners of expression and develops their emotions within the school environment (Zembylas, 2006). We should also consider that their lives transcend the physical barriers of educational centres and bear in mind their interaction in the school environment with their representations of their everyday selves as parents, friends, women/men, etc.

To conclude, this paper briefly describes the most important points that should be taking into consideration in research into the understanding of emotional phenomena of education professionals in "special" circumstances.

## References

- Allan, J. (1999) *Actively seeking inclusion: pupils with special needs in mainstream schools*, London: Falmer Press
- Atkinson, P., Coffey, A., Delamont, S. et al (Eds) (2001) *Handbook of ethnography*. London: Sage
- Boler, M. (1999) *Feeling power: emotions and education* NY: Routledge
- Bone, D (2002) Dilemmas of emotion work in nursing under market-driven health care. *Intentional journal of public sector management*, 15 (2) 140-150).
- Bradberry, T; and Greaves, J. (2005). *The Emotional Intelligence. Quick Book*. New York: Simon and Schuster).
- Butler, J (2005) *Giving an account of oneself*, Folham University Press
- Cannon, W. B. (1927) *Bodily Changes In Pain, Hunger, Fear And Rage - An Account Of Recent Researches Into The Function Of Emotional Excitement*. London: Cannon Press
- Cheek, J; and Gough, N. (2005) "Postmodernist perspectives", in Somekh, B and Lewin, C. (Eds.) *Research Methods in the Social Sciences*. London: Sage.
- Clarke, A. (2005) *Situational analysis: grounded theory after the postmodern turn*, London: Sage
- Coffey A (2002) Ethnography and self: reflections and representations, in T May (ed) *Qualitative Research in Action* (London: Sage) pp.313-31.
- Crespo, E. (1986) "Chap 11 Regional variation: emotions in Spain" in R.Harre, *The social construction of emotions*. Basic Blackwell.
- Croysant, J and Restivo, S. (1995) Science, Social Problems, and Progressive Thoughts: Essays on the tyranny of science. pp 39-88. In Star, S. L. (ed) (1995) *Ecologies of knowledge : work and politics in science and technology*. Albany : State University of New York Press
- Damasio, A.R. (1995) *Descartes' error : emotion, reason and the human brain*. London: Picador
- Derrida, J. (1976) *Of grammatology*, Baltimore and London: Johns Hopkins University
- Drolet, M. (2004) *The postmodernism reader : foundational texts*. London: Routledge
- Edwards, D. (1999) Emotion Discourse, in *Culture & Psychology*; vol. 5: pp. 271 – 291
- Edwards, D.; Ashmore, M.; and Potter, J. (1995) "Death and furniture: the rhetoric, politics and theology of the bottom line arguments against relativism" *History of the human sciences* 18 (2) pp. 25-49
- Ellis, C and Bochner, A. (eds) (1996) *composing ethnography: Alternative forms of qualitative writing*. Walnut Creek, CA: AltaMira.
- Feyerabend, P. (1975) *Against method: outline of an anarchistic theory of knowledge*. London: NLB

- Fine, M. and Weis, L. (1998) "Writing the "wrongs" of fieldwork: confronting our own research /writing dilemmas in urban ethnographies", In G. Shacklock and J. Smyth (eds), *Being Reflexive in Critical Education and Social Research*. London: Falmer Press, pp 13-35
- Foucault, M. (1989) "Archaeology of Knowledge". London; New York: Routledge
- Foucault, M. (1995) "Power/Knowledge" in Rabinow, P. *The Foucault Reader*, London: Penguin
- Foucault, M. (1995b) "Truth and Power" in Rabinow, P. *The Foucault Reader*, London: Penguin
- Foucault, M. (2000) "What is an author" in D. Faubion (eds), *Essential works of Foucault 1954-1984 : aesthetics Volumen 2*. London: Penguin
- Gardner, H (1989) "Multiples intelligences go to school" in *Educational Research* 18 , 8
- Gardner, H. (1995) *Inteligencias multiples. La teoria en la practica*. Barcelona: Pidos
- Gee, J.P (1999) *An introduction to Discourse Analysis: Theory and Method*, London: Routledge
- Goldie, P (2000) *The emotions: A philosophical exploration*. Oxford: Oxford University Press
- Goleman, D. (1996) *Emotional intelligence : why it can matter more than IQ*. London: Bloomsbury
- Griffiths, P.E. (1997) *What Emotions Really Are*, Chicago: University of Chicago
- Hargreaves, A. (1998) *Changing Teachers, Changing times: Teachers work and culture in the postmodern age* ,London: Redwood
- Hargreaves, A; Moore, Manning (2001) *Learning to change: subjects beyond subjects and standards*, San Francisco: Willey
- Harre, R. (Ed.)(1986) *The social construction of emotions*. Basil: Blackwell
- Harre, R and Stearn, P. (1995) *Discursive psychology in practice*. London: Sage
- Hochschild, A. R. (2003) *The managed heart; commercialization of human feeling* 20<sup>th</sup> anniversary edition. Los Angeles: University of California press
- Hughes, C (2002) *Key concepts in feminism theory and research*, London: Sage.
- James, W (1984) What is an emotion? *Mind* 9, 188-205
- Ladson-Billing, G. (1998) "Just what is critical race theory and what's doing in a nice field like education?". *Qualitative Studies in Education*, 11(1) 7-24.
- Lather, P. (1993) Fertile obsession: validity after poststructuralism, *Sociological Quarterly*.34 (4)
- Le Doux, J. (1998) *The emotional Brain*. London: Orion
- Liotard, J (1984) *The postmodern condition: a report on knowledge*, Manchester: Manchester University Press
- MacLure, M. (2003) *Discourse in Education and Social Research*, Midenhead: Open University Press.
- Mortensen, C. (2000) Plato's pharmacy and Derrida drug store, in *Language & Communication*, v20 n4 p329-46 Oct 2000
- Parker, L.; and Roberts, L. (2005) " Critical theories of race" in Somekh, B and Lewin, C. (Eds.) *Research Methods in the Social Sciences*. London: Sage.
- Pastor, J.; Ovejero, A. (2007) *Michael Foucault: Caja de Herramientas contra la dominación* [Michael Foucault: box of tools against domination] Oviedo: Universidad Oviedo
- Plumwood, V. (1993) *Feminism and the Mastery of Nature*, London:Routledge
- Roth, P. A (1987) *Meaning and method in the social sciences : a case for methodological plural* Ithaca; London : Cornell University Press
- Salovey, P. and Mayer, J.D. (1990). Emotional Intelligence. *Imagination, cognition and personality* , 9, 185-211
- Smith, D.E. (1987) *The Every Day World as Problematic: A Feminist Sociology*, Boston: Northeastern University Press.
- Stearns, C.Z. and Stearns, P.N. (Ed.) (1998) *Emotion and Social Change: Toward a New Psychohistory*. London: Holmes & Meier
- Jencks, C. (1995) *A post-modern reader*. St. Martins Press
- Wetherell, M; Taylor, S; Yates, S. (2001) *Discourse Theory and Practice: a reader*, London: Sage
- Woolgar, S (1988) *Science: the very idea*. London: Routledge
- Zembylas, M. (2005) Discursive Practice, genealogies and emotional rules: a poststructuralist view on emotion and identity in teaching. *Teaching and teacher education*, 21 935-948).
- Zembylas, M. (2006) *Challenges and possibilities in a postmodern culture of emotions in education* . *Interchanges* 37/3 pp. 251-275.

# Comparison of ADMS 4 and LiDAR in the Prediction of Atmospheric Boundary Layer

**Vera Bachtiar, Chris Collier, Fay Davies**, *Centre of Environmental Research System, School of Environment and Life Sciences, University of Salford*

## Abstract

Since the late 1950s, atmospheric dispersion models have been developed to predict air quality. The Atmospheric Dispersion Modelling System (ADMS) is a dispersion model that simulates buoyant and neutrally buoyant plumes. Here the model is used to compare modelled and measured boundary layer depth in a city centre where traditionally models are not reliable. Doppler LiDAR (Light Detection and Ranging) can measure basic boundary layer variables, such as mixing depth, wind shear and turbulence, and inversion height. Comparisons are made between ADMS 4 and the Salford LiDAR for data taken in central London on 29 October 2007. On average, boundary layer height predicted by the LiDAR is higher than ADMS 4. The ADMS model has a very simple surface scheme that is not representative of a complex urban environment. The results show that there is not sufficient surface roughness within the model to produce a high enough boundary layer depth. The final aim of this research is to produce a modification to the model to enable the correct simulation of this complex terrain.

**Keywords:** Boundary layer height, ADMS 4, Doppler LiDAR

## 1. Introduction

Since the late 1950s, atmospheric dispersion models have been developed to predict air quality (Carruthers, 1994). More than 20 years later, advanced dispersion models have been produced. An example is the Atmospheric Dispersion Modelling System (ADMS). ADMS is used for pollution dispersion modelling in several countries in Europe, Asia, Australia, and the Middle East (CERC, 2008).

In the UK, ADMS was used to review air quality in central London in 1996/1997 and assess future air quality against air quality objectives in 2005 (Colvile et al., 2002). According to CERC (2008), users of ADMS in the UK include over 130 individual company licence holders and regulatory authorities. Amongst them are the UK Health and Safety Executive (HSE), the Environment Agency in England and Wales, the Scottish Environmental Protection Agency (SEPA), the Environment and Heritage Service in Northern Ireland, and government organisations such as the Food Standard Agency.

One of the capabilities of ADMS 4 is the prediction of atmospheric boundary layer properties—most importantly, boundary layer height (BLH). According to Stull (1988, p.2), boundary layer is “that part of the troposphere that is directly influenced by the presence of the earth’s surface and responds to surface forcing with a timescale of about an hour or less.”

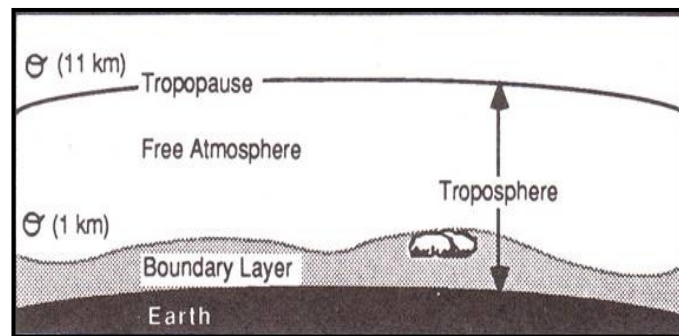
BLH is one parameter of the meteorological data that is required in the meteorological pre-processor in the ADMS model. It is an important parameter in dispersion models because determines the location and transport of aerosols and pollutants. BLH also will determine the volume available for pollutant dispersion, depending on meteorological parameters, surface turbulent fluxes, and physical parameters (Fisher et al., 2005). The limits on the vertical diffusion of the plume or puff of material released are also determined by BLH. In urban air quality modelling, BLH is influenced by the vertical profile of mean wind velocity and the turbulent vertical exchange of momentum, heat, and moisture (Dandao et al., 2009).

The accuracy of BLH is important in urban air quality modelling because it affects near-surface pollutant concentrations (Dandao et al., 2009). BLH estimation is different in different models. An example of different BLHs calculated by ADMS 3.1 and AERMOD PRIME 02222 was shown by Sidle et al. (2004). They found that the pollution concentration differences between the two models could be attributed to the differences in the derived BLHs. In this paper, we carried out prediction of the atmospheric boundary layer (BLH) depth using the ADMS 4 model and compared it to measurements from LiDAR aerosol backscatter data.



## 2. Atmospheric boundary layer

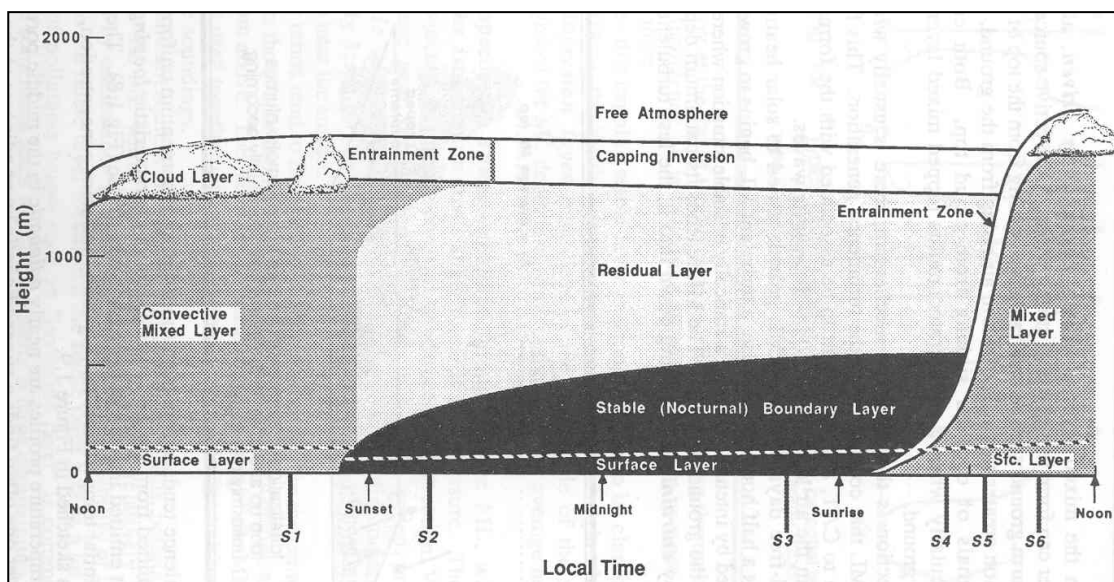
The atmospheric boundary layer is part of the troposphere. The troposphere consists of a boundary layer and free atmosphere, as shown in Figure 1. The boundary layer's thickness is variable from as low as 100 metres to 2–3 kilometres. This depends on several factors, including evaporation and transpiration, frictional drag, heat transfer, pollutant emission, and terrain-induced flow modification. The height of boundary layer is not constant with time and depends on the strength of the surface-generated mixing. In the daytime, when the earth's surface is heated by the sun, there is an upward transfer of heat from the surface to the atmosphere. In contrast, at night-time, the Earth's surface cools more rapidly than the atmosphere (Oke, 1987).



**Figure 1.** Troposphere and atmospheric boundary layer.

Source: Stull (1988)

The diurnal cycle depth of the boundary layer over land surfaces can be seen in Figure 2. The major components of the structure are the mixed layer, the residual layer, and the stable boundary layer.



**Figure 2.** Structure of the atmospheric boundary layer.

Source: Stull (1988)

On a cloud-free day, after about one hour after sunrise, thermally driven mixing begins. In this condition, thermals of warm air rise from the ground. As the surface heats up through the morning, this well-mixed layer grows in height, causing an increase in the depth of the statically unstable layer. In the late afternoon, the mixed layer reaches its maximum depth. Because of the turbulence, the characteristics of the mixed-layer heat, moisture, and momentum profiles are constant in the vertical direction. In the middle of the mixed layer,

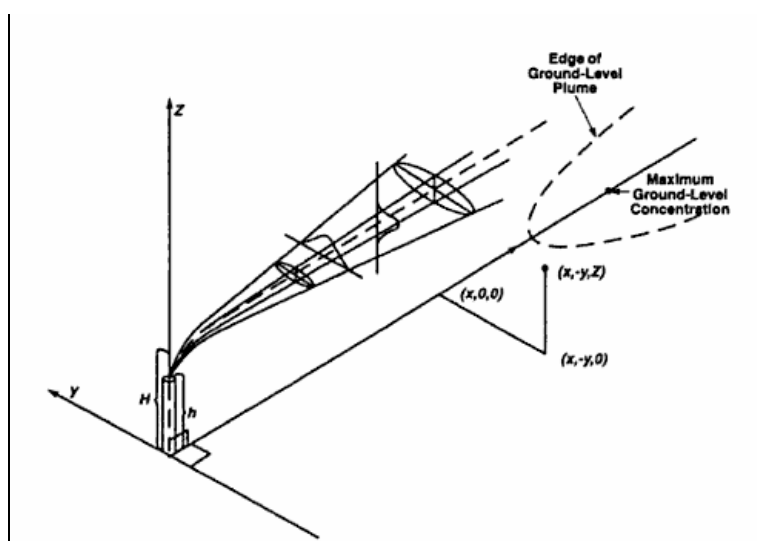
virtual temperature profiles are nearly adiabatic, but in the surface layer they tend to be super-adiabatic. At the top of the mixed layer, a stable layer is present that limits the height to which the thermals rise. The height of this level is defined as the BLH. The zone between the well-mixed layer and the free troposphere is called the entrainment zone, because in this zone entrainment into the mixed layer occurs (Stull, 1988).

About one hour before sunset, turbulence decays in the previous well-mixed layer. The layer has pollution concentrations similar to those of the previous well-mixed layer. This layer is called the residual layer and has neutral turbulence and equal intensity in all directions (Stull, 1988).

In the night-time, the bottom portion of the residual layer transforms into a stable boundary layer by its contact with the ground. This layer is characterised by statically stable air with weaker and sporadic turbulence. Although surface winds are calm at night, the winds aloft increase rapidly and form a low-level jet or nocturnal jet. This nocturnal jet tends to generate turbulence, so that sometimes turbulence occurs in the stable boundary layer (Stull, 1988).

### 3. Atmospheric dispersion models

Atmospheric dispersion models are computer programs using mathematical algorithms to simulate pollution dispersion in the atmosphere. They have been developed since the late 1950s. Most dispersion models use a Gaussian dispersion process and meteorological surface measurement (Carruthers, 1994). An illustration of Gaussian dispersion process (the Gaussian Plume Model) is provided in Figure 3.



**Figure 3.** Gaussian Plume Model.

Source: Water Environment Federation, American Society of Civil Engineers, Water Environment Federation. Liaison Subcommittee (1995)

Gaussian models are based on a Gaussian distribution of the plume in the vertical and horizontal directions under steady-state conditions. The normal distribution of the plume is modified at greater distances due to the effects of turbulent reflection from the surface of the earth and at the boundary layer when the mixing height is low (Holmes and Morawska, 2006). The width of the plume is determined by the standard deviation of the concentration distribution in the crosswind direction ( $\sigma_y$ ) and the standard deviation of the concentration distribution in the vertical direction ( $\sigma_z$ ) (Turner, 1994).

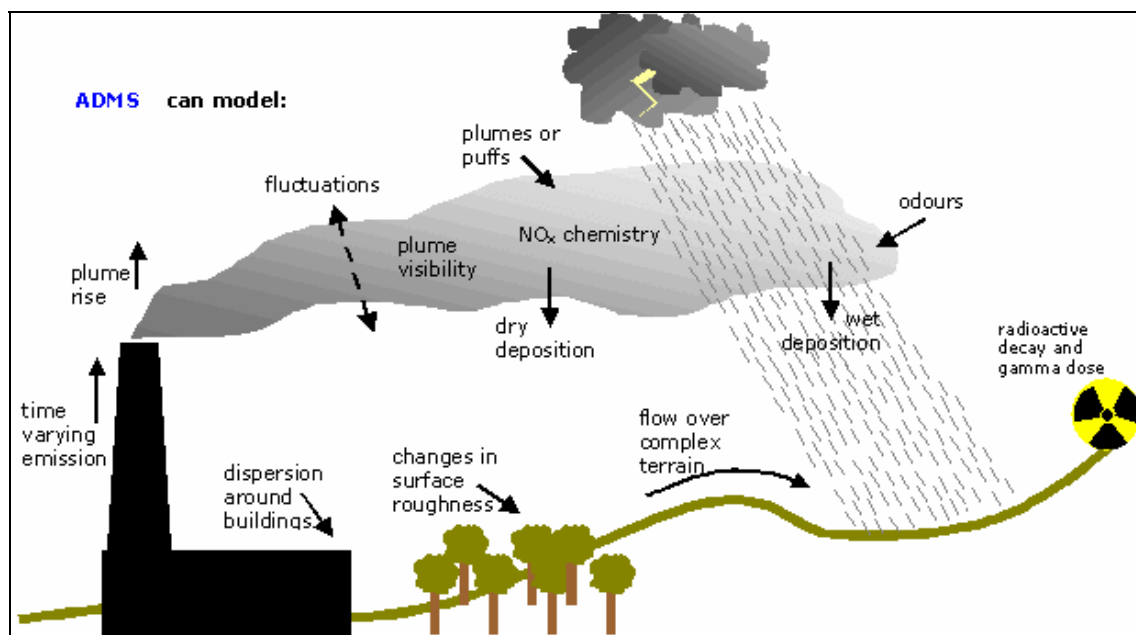
Over the past 20 years, better and better models have been developed. Current models use a new approach—namely, the vertical profile of mean velocity, temperature, and turbulence in the boundary layer above the ground; the height of the boundary layer  $h$ ; and the Monin-Obukhov length  $L_{MO}$  (Carruthers, 1994).

Atmospheric dispersion models have been developed in the United States, the United Kingdom, continental Europe, and Australia. The models range from screening models to advanced models. The U.S. Environmental Protection Agency's (EPA) list of recommended models includes the AERMIC MODEL (AERMOD), CALPUFF, BLP, CALINE3, CAL3QHC and CAL3QHCR, CTDMPPLUS, and OCD. Other models include the Air Force Dispersion Assessment Model (ADAM), the Atmospheric Dispersion Modelling System (ADMS), Industrial Source Complex (ISC3), and the Air Force Toxics Model (AFTOX). Some of the models that are used in the United Kingdom are GASTAR, NAME, UDM, ADMS-screen, ADMS 3, ADMS-URBAN, ADMS-roads, and ADMS 4 (the newest).

#### 4. ADMS 4

Atmospheric Dispersion Modelling System (ADMS) is a dispersion model that simulates buoyant and neutrally buoyant particles and gasses (Carruthers et al., 1994). The model can predict the boundary layer structure. It uses normal Gaussian distributions in a stable and neutral condition (Holmes and Morawska, 2006). ADMS was developed by a government and industry consortium in the UK (CERC, 1998; in Hana et al., 2001).

As illustrated in Figure 4, ADMS can simulate a number of processes. The figure shows a number of model processes. Firstly, ADMS 4 can model dry and wet deposition. It can assume dry deposition to the near-surface concentration. Wet deposition is modelled through a washout coefficient, irreversible uptake is assumed, and plume strength following wet deposition decreases with downwind distance. Secondly, ADMS 4 also can model continuous (i.e., plumes and time-dependent) release (i.e., puff). Thirdly, ADMS 4 can model the variation of the emission rate with time, building effect, complex terrain, and coastlines. The model can include up to 25 buildings in each run. For running coastlines, the model assumes that the sea is lower than the land, that there are convective meteorological conditions on the land, and that there is onshore wind. Furthermore, ADMS 4 can model the dispersion of odour and predicts the decay of radioactive (and gamma dose) species released from a source. Finally, ADMS 4 can model a simple  $\text{NO}_x$  chemistry scheme involving the conversion of Nitrogen Dioxide ( $\text{NO}_2$ ) to Nitrous Oxide ( $\text{NO}$ ) and Ozone ( $\text{O}_3$ ).



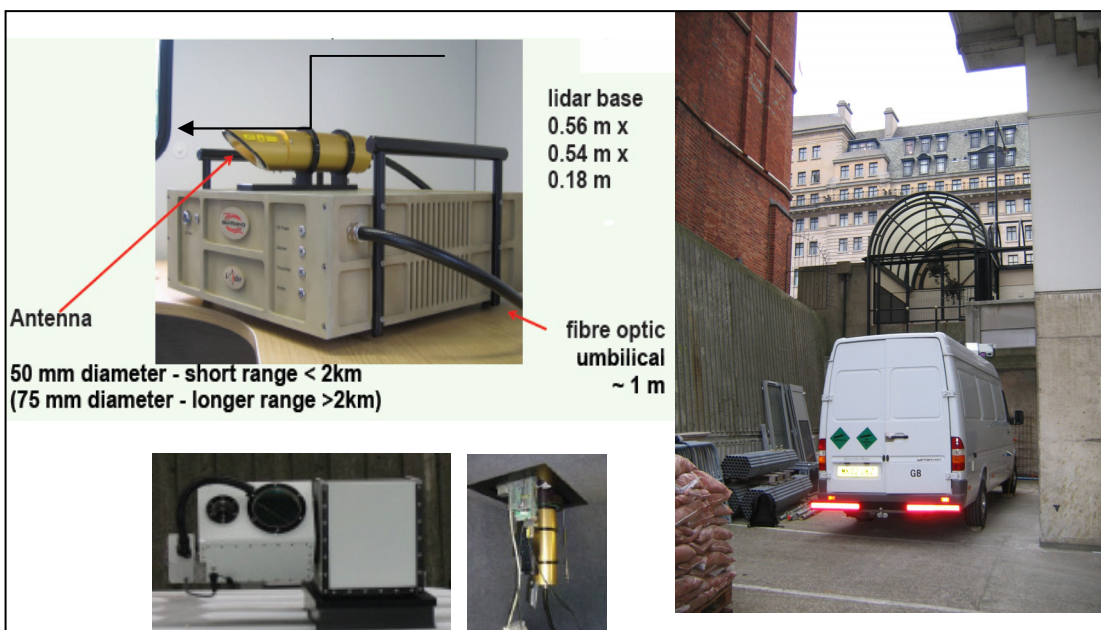
**Figure 4.** Processes that can be represented by ADMS.

Source: CERC (2008)

## 5. Salford Doppler LiDAR

Doppler LiDAR (Light Detection and Ranging) is an instrument that transmits pulses of light from a laser to the atmosphere, which are then reflected from aerosols suspended in the atmosphere. This weakly reflected signal is collected by telescope. The distance from which the reflection occurs is calculated by relating the speed of light to the timing of the transmitted and received pulses (Koch, 2009). A pulse or pulses of light are transmitted by the LiDAR system, and then the backscatter signal or intensity of the return signal is received back at the receiver. The backscatter signal or intensity is analysed by the LiDAR system as a function of time. A LiDAR can measure basic boundary layer variables, such as aerosol backscatter coefficient and depolarisation ratio, longitudinal (along the beam) wind component, water vapour density, temperature, and concentrations of some other constituents. From these basic measurements, other boundary layer parameters can be calculated, such as mixing depth, wind shear, inversion height, and aerosol type (Schwielsow, 1986).

The Salford Halo Doppler LiDAR (Figure 5) is an autonomous instrument for atmospheric remote sensing that operates at a wavelength of 1.5 microns. It employs novel optical technology, is designed to be eye-safe class 1M, and has low power. The system has three separate units: the optical base unit, the weatherproof monostatic/antenna, and the signal-processing and data-acquisition unit. The optical base unit has dimensions of 0.56 m × 0.54 m × 0.18 m, and contains the optical source, interferometer, receiver, and electronics. The weatherproof antenna is connected to base unit using an umbilical. The antenna can be placed in outside, whilst the base unit and data-acquisition system are housed within a laboratory environment or environmental container (Bozier et al., 2007).



**Figure 5.** Salford Halo Doppler LiDAR.

Source: Davies et al. (2008)

The system parameters of the Salford Halo Doppler LiDAR can be seen in Table 1. The parameters—such as range gate, maximum range, number of pulses accumulated for each measurement, and the temporal resolution of the Doppler measurements—can be set by the user. The LiDAR system can be monitored and controlled by remote access software that transfers data using an internet network connection (Bozier et al., 2007).

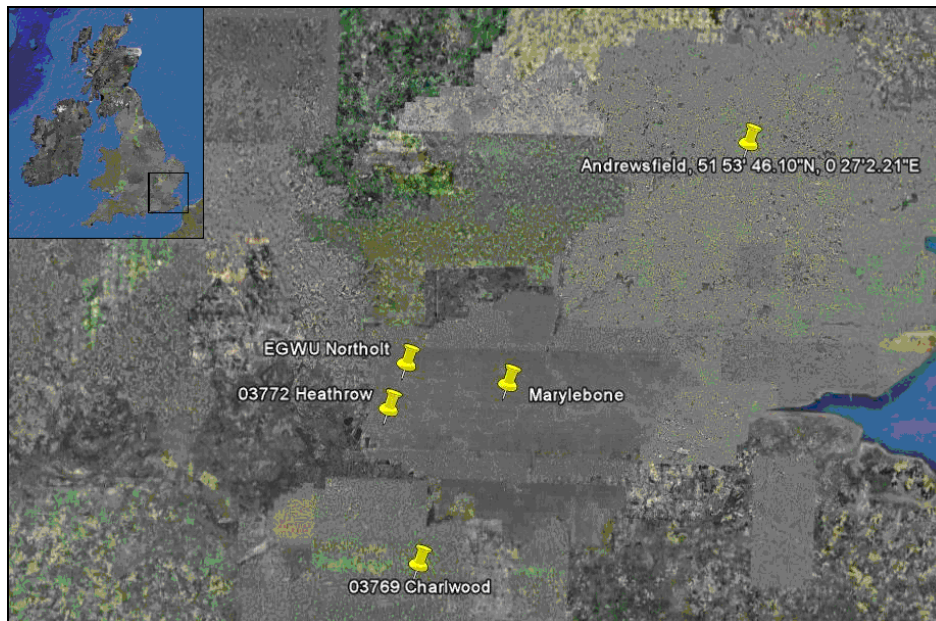
**Table 1.** The Salford Halo Doppler LiDAR parameters.

Parameter	Value
Operating wavelength	1.55 $\mu\text{m}$
Pulse repetition frequency	20 kHz
Pulse duration	150 ns
Beam divergence	50 $\mu\text{rad}$
Range gate	Variable: 20-60 m
Minimum range	$\sim$ 50 m
Maximum range	7 km
Temporal resolution	0.1-30 s

Source: Davies et al. (2008)

## 6. Results and discussion

Comparison was made between ADMS 4 and Salford LiDAR on 29 October 2007. The input surface meteorological data for ADMS 4 were obtained from the Met Office. Because no surface meteorological data is available from central London, data was used from several locations close to where the LiDAR data was measured. These locations are Andrewsfield, Charlwood, Heathrow, and Northolt, as shown in Figure 6. The LiDAR data was taken by the University of Salford pulsed Doppler LiDAR, which was based at the University of Westminster on Marylebone Road in London. These locations can be seen in Figure 6.

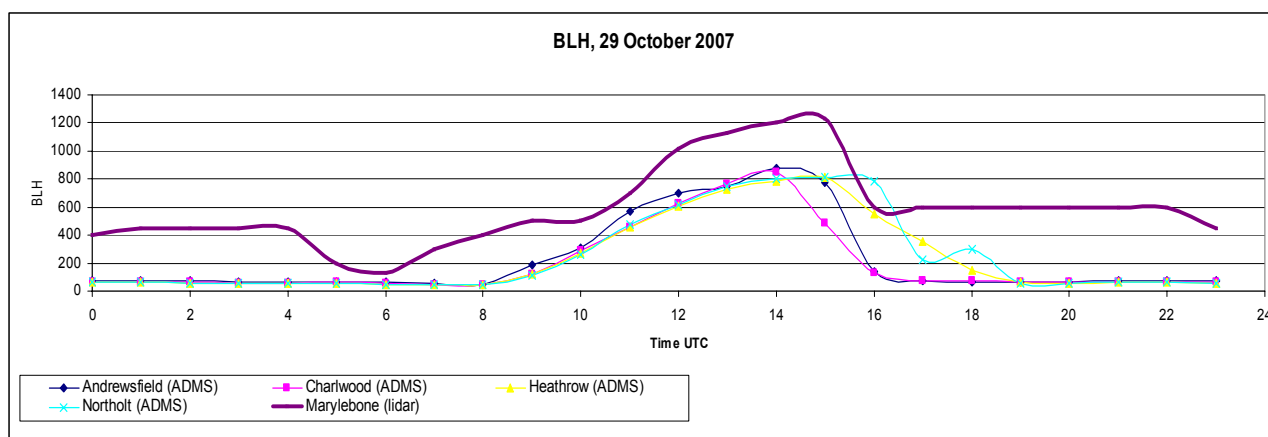


**Figure 6.** Location of study.

Source: Google Earth (2009)

The boundary layer heights predicted by ADMS and measured by the LiDAR on 29 October 2007 can be seen in Figure 7. This figure shows that boundary layer heights measured by the

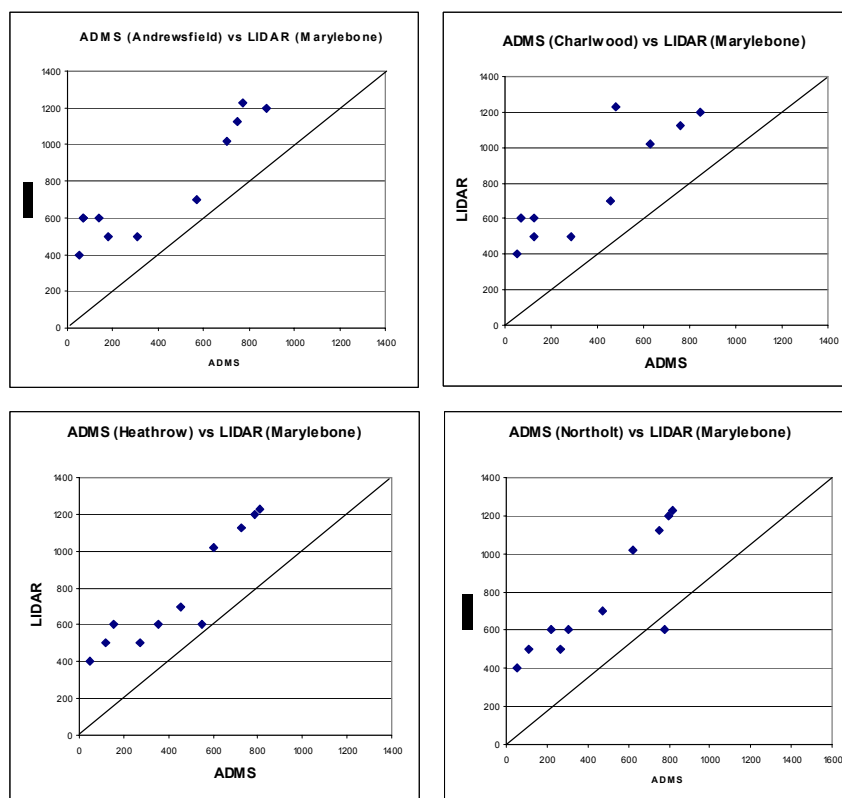
LiDAR are higher than those predicted by ADMS 4. In the daytime, in well-mixed conditions, the highest boundary layer heights shown by the LiDAR data was 1230 m, and ADMS predicted the highest boundary layer height to be 846 m. During the night-time, there is an assumption within ADMS that the boundary layer collapses to a stable nocturnal layer. In these cases, ADMS assumes a constant MHL of 100 m, as shown in Figure 7. In urban areas, however, the night-time boundary layer can still contain mixing driven by anthropogenic heating and turbulent mixing of the flow around the buildings. The LiDAR MLH measurements are affected by both of these factors, but are also determined by the aerosol in the atmosphere, which can persist even without mixing in the residual layer. For this reason, the night-time predictions of the BLH are too low from ADMS and if a residual layer is present, the measurements of the LiDAR BLH could be too high. This phenomenon can be seen after 16 UTC, when the boundary layer height decreases in ADMS while LiDAR prediction remains constant through most of the evening.



**Figure 7.** Boundary layer height (BLH) on 29 October 2007.

Correlations between ADMS 4 and LiDAR in the daytime are shown in scatter plots in Figure 8. The figure shows scatter plots of ADMS 4 predictions (for Andrewsfield, Charlwood, Heathrow, and Northolt) against LiDAR data (in Marylebone). The figure shows large amounts of scatter. The 1:1 correlation between ADMS 4 and LiDAR data is shown by the 45° lines in the figure. These showed that ADMS 4 values in all locations are equal to about half those of the LiDAR data values.

The difference in BLH from ADMS and LiDAR can be due to different factors. Firstly, meteorological input used in ADMS 4, such as wind speed, was from a meteorological station outside of the city, where there was lower surface roughness (Davies et. al., 2007). Furthermore, ADMS 4 has a very simple surface scheme, which is not representative of a complex urban environment. The results show that there is not sufficient surface roughness within the model to produce a high enough boundary layer depth.



**Figure 8.** Correlation between ADMS 4 and LiDAR data in the daytime.

A comparative statistical analysis has been carried out between ADMS 4 and LiDAR. Table 2 shows coefficient of determination ( $R^2$ ) and percentage difference between ADMS 4 and LiDAR. The coefficient of determination represents the percent of the data that is the closest to the line of best fit. These data shows  $R^2$  values between 0.70 and 0.85, with an average value of 0.79. Meanwhile, percentage difference between ADMS 4 and LiDAR is between 43% and 60%, with an average of 51%. This means that ADMS 4 and LiDAR data have good correlations in pattern, but both have large differences in value—namely, all ADMS 4 data is underestimated, in contrast to LiDAR data.

**Table 2.** Statistical analysis.

ADMS locations	$R^2$ (Coefficient of determination)	% difference with LiDAR
Andrewsfield	0.85	54 %
Charlwood	0.77	60 %
Heathrow	0.85	47 %
Northolt	0.70	43 %
average	<b>0.79</b>	<b>51 %</b>

## 7. Conclusion

Comparison between ADMS 4 and LiDAR data was carried out on central London data on 29 October 2007. Meteorological data for ADMS 4 was obtained from the Met Office for four meteorological stations: Andrewsfield, Charlwood, Heathrow, and Northolt. LiDAR Data was taken by the University of Salford pulsed Doppler LiDAR, which was based at the University of Westminster on Marylebone Road in London.

The comparison shows that the boundary layer height (BLH) predicted by the LiDAR is higher than that predicated by ADMS 4. There is a lot of scatter in the data. This can be due many different factors, including meteorological variables and surface roughness. Finally, ADMS 4 has a very simple surface scheme that is not representative of a complex urban environment. The results show that there is not sufficient surface roughness within the model to produce a high enough boundary layer depth.

Further work will aim to develop a surface model to improve the estimation of BLH from ADMS and thereby enable the correct simulation of this complex urban environment.

## Acknowledgments

This study is funded by the Indonesian government's Directorate General of Higher Education. We also acknowledge the Met Office of the United Kingdom for providing us with the meteorological data for running ADMS 4.

## References

- Bozier, K.E. et al. (2007) 'Evaluation of a new autonomous Doppler LiDAR system during the Helsinki International Testbed Field Campaign,' *Third Symposium on LiDAR Atmospheric Applications, American Meteorology Society Annual Meeting, San Antonio, Texas, 15-18 January 2007*; available at: <http://ams.confex.com>.
- Carruthers, D.J. et al. (1994) 'UK-ADMS: a new approach to modelling dispersion in the earth's atmospheric boundary layer,' *Journal of Wind Engineering and Industrial Aerodynamics* 52 (1-3), pp. 139-153.
- Cambridge Environmental Research Consultants (CERC) (2008) 'ADMS 4 industrial air pollution model'; available at: <http://www.cerc.co.uk>.
- Colville, R.N. et al. (2002) 'Uncertainty in dispersion modelling and urban air quality mapping,' *Environmental Science & Policy* 5 (3), pp. 207-220.
- Dandao, A. et al., (2009) 'A comparison between modelled and measured mixing-layer height over Munich,' *Boundary-Layer Meteorology* 131 (3), pp. 425-440.
- Davies, F. et al. (2008) 'The Salford Halo Pulse Doppler LiDAR System,' *National Centre Atmospheric Science Annual Conference 2008*; available at: <http://www.ncas.ac.uk>.
- Davies, F. et al. (2007) 'Urban air pollution modelling and measurements of boundary layer height,' *Atmospheric Environment* 41 (19), pp. 4040-4049.
- Fisher, B. et al., eds. (2005) *Meteorology Applied to Urban Air Pollution Problems: Final Report of COST Action 715*. Demetra Ltd Publishers.
- Hana, S.R. et al. (2001) 'Evaluation of the ADMS, Aermot, and ISC3 dispersion models with the Optex, Duke Forest, Kincaid, Indianapolis, and Lovett field datasets,' *International Journal of Environment and Pollution* 16 (1-6), pp. 301-314.
- Holmes, N.S. and Morawska. L. (2006) 'A review of dispersion modelling and its application to the dispersion of particles: an overview of different dispersion models available,' *Atmospheric Environment* 40 (30), pp. 5902-5928.
- Koch, G. (2009) 'Doppler Lidar observations of an atmospheric thermal providing lift to soaring ospreys,' *NASA Langley Research Centre, Hampton, VA*; available at: <http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov>.
- Oke, T.R. (1987) *Boundary Layer Climates*. London and New York: Routledge.
- Schwiesow, R.L. (1986) 'Probing the atmospheric boundary layer,' *American Meteorological Society Boston, MA*, pp.139-162.



- Sidle, C. et al. (2004) 'A comparison of boundary layer heights in atmospheric dispersion models and possible consequences for modelling emissions from tall stacks,' *Ninth International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes*, Garmisch-Partenkirchen, Germany.
- Stull, R.B. (1988) *An Introduction to Boundary Layer Meteorology*. Dordrecht: Kluwer Academic Publishers.
- Turner, D.B. (1994) *Workbook of Atmospheric Dispersion Estimates: An Introduction to Dispersion Modeling*. Boca Raton, FL: Lewis Publishers.
- Water Environment Federation, American Society of Civil Engineers, Water Environment Federation. Liaison Subcommittee (1995) 'Odor control in wastewater treatment plants; available at: <http://www.cerc.co.uk/software/adms4.htm>.

# Improving Integration in the Construction Industry through Knowledge Management in a Relational Contracting Setting: A Critical Success Factors Approach

**Anis Sazira Bakri, Bingunath Ingirige, and Dilanthi Amaratunga**, *School of the Built Environment, University of Salford*

## Abstract

Construction is a project-based industry. In every project, there is generally a unique product, requiring multiple reconfigurations. This means that construction supply chains are characterised by adversarial practices and fragmentation in relationships. As a result, construction project participants generally have temporary relationships rather than long-term partnerships. A consequence of this is lack of trust amongst construction clients, designers, main contractors, and subcontractors. Since the supply chain works as a disparate collection of separate organisations rather than as a unified team, it suffers from a lack of integration. Besides this, construction participants are often unwilling to share knowledge, resulting in poor knowledge flow. These characteristics of the construction industry are the main reasons for its low efficiency and productivity in project delivery. It is therefore important to focus on the management of knowledge within the construction industry setting. Knowledge Management (KM) is aimed at improving collaborative behaviour between firms and individuals involved in the construction process, such as knowledge sharing, knowledge transfer, and communications and information technology. In addition, Relational Contracting (RC) is one way to minimise the problems of fragmentation. It is believed that RC, when integrated with KM, can successfully address the major problems of the industry and its clients. There are some common key issues shared between RC and KM. Therefore, the aim of this paper is to discuss how KM through its critical success factors (CSFs) could be used to improve integration in RC project settings.

**Keywords:** Integration, relational contracting, knowledge management, critical success factors

## 1.0. Introduction

From the pre-design to the design stages of projects, the construction industry involves a wide range of activities and participants. This industry is often characterised as fragmented when the traditional construction process involves players that are disconnected and isolated from each other, resulting in inefficiencies (Latham, 1994). Each activity in the industry generates and uses information; the different participants need to engage in an effective communication network so that relevant information and knowledge flows are not impeded. Poor communication, information transmission, coordination, and teamwork have been identified as the causes of most of the performance problems in the construction industry (Kajewski et al. (2003). As a result of the lack of integration in the industry (Akintoye et al., 2000), the traditional construction process tends to incur additional costs for rework stemming from quality issues, disputes, and slower building times.

A report by Constructing Excellence (2004) indicates that successful projects result from a combination of effective planning, team collaboration and integration, maintenance of knowledge, and flexibility in design during the construction process. The Latham report (1994) and the Egan report (1998) suggest that the industry could achieve better performance through greater teamwork amongst the construction participants. Recommendations within these reports have led to an increased use of inter-organisational relationships in order to improve the construction development process. These include the relational contracting/relationship contract (RC) principles that underpin approaches such as partnering, strategic alliances, joint venture, public-private partnerships, and supply chain management. However, it may be anticipated that not all the collaborative relationships in construction

developments will be successful. Having experienced the same problems, the Malaysian construction industry is attempting to improve the situation through the development of the Construction Industry Master Plan (CIMP) 2006–2015. Seven strategic initiatives have been identified, and these have subsequently formed the basis for the CIMP's main recommendations. One of these is to integrate the construction industry value chain in order to enhance productivity and efficiency and to minimise the problem of fragmentation. Collaborative approaches such as partnering are also a component of this initiative.

The importance of knowledge in different types of inter-organisational relationships has been acknowledged in the literature (Khamseh and Jolly, 2008). Knowledge Management (KM) is one of the important ways for organisations to identify, transfer, and internalise external knowledge. Kogut (1988) notes that one of an organisation's main motivations for entering into a collaborative arrangement is to transfer organisational knowledge. The diverse settings brought under the umbrella of Relational Contracting (RC)—such as partnering, joint ventures, alliances (Ingirige and Sexton, 2006) and supply chain management—have created unique knowledge sharing and learning opportunities for the partner organisations (Naphiet and Ghoshal, 2005). To assure their success, more strategic approaches to managing construction projects are needed. KM and Information Technology (IT) infrastructure are seen as vital in this regard (Wiig, 1997), as they could minimise wasteful activities and improve productivity and efficiency. Although the importance of KM has been widely promoted and recognised, it seems that few organisations are truly capable of leveraging and managing knowledge in their organisations. According to Storey and Barnett (2000), a significant proportion of KM initiatives have failed due to the complexity of the different type of knowledge to be managed, which involves the support of a technological infrastructure and a change in organisational culture. In order to achieve effectiveness for KM implementation, it is important to identify and examine its critical success factors (CSFs) and to understand how these factors can contribute to the success of RC for better integration in the construction industry. RC is essential to the functioning of KM. It creates collaborative relationships throughout the construction supply chain and highlights the importance of trust, communication, commitment, cooperation, and joint problem solving. KM, together with RC, ensures that knowledge and information are shared with the construction partners. Previous studies have suggested several generic CSFs for KM related to people, processes, and technology.

In the Malaysian context, existing research in KM is confined to examining its role and relationship to organisational factors (Narayanan et al., 2003). The extent to which KM is practised in construction industry has yet to be established. The ways in which construction organisations (clients, consultants, contractors, etc) currently manage their knowledge can be explored in order to propose best practices. Several studies (Mohamed et. al., 2007; Bakri and Saidan-Khaderi, 2007) have identified the perceptions and understanding of KM in construction organisations, but there are still insufficient studies to identify the main drivers or enablers and success factors of KM in this industry. These are needed in order to better understand the major obstacles in knowledge sharing and how to integrate KM for different projects and contractual approaches in the construction processes. In the previous literature, little effort has been made to identify and examine the factors that are critical for successful implementation of KM in the context of RC. It is important to investigate the roles of KM in order to improve project performance and integration, and to understand the barriers to implementing KM within RC projects. This research addresses this gap in the literature and recommends guidelines for construction participants to improve integrations within RC projects by considering critical project-orientated CSFs.

## **2.0. Aim of the research**

As noted, this research aims to develop suitable guidelines for improving integration in the construction industry. To this end, it considers critical project-oriented CSFs for KM implementation within a relational contracting project setting. Since KM is aimed at improving construction management performance and assuring client satisfaction, it requires the understanding of the factors that are critical for the success of the KM implementation. To achieve this aim, the present research seeks to (1) identify the roles of KM and the barriers to improving integration within RC project settings in the Malaysian construction industry, and (2) identify and rank the CSFs for KM implementation within RC project settings.

### 3.0. Literature review

#### 3.1. Fragmentation and integration in construction industry

The construction industry is often characterised as fragmented when the traditional construction process involves players that are disconnected and isolated from each other, resulting in inefficiencies (Egan, 1998). Because a large number of construction firms are geographically dispersed (Latham, 1994), project teams often have to work together virtually, from many different locations over the life of a project. As a result of the lack of integration in the industry, the traditional construction process tends to incur additional costs for rework stemming from quality issues, disputes, and slower building times. Therefore, in addition to the immediate project team, construction firms need to further manage networks of suppliers, customers, and regulatory bodies, as knowledge is differentiated and distributed within, and across, these supply networks (Gann and Salter, 1998).

The problem of fragmentation can significantly hinder learning and knowledge advancement within the industry (Senaratne, 2005). At a project level, knowledge and learning is very much affected by the activities of project team members. Therefore, the manner in which the construction project team members are integrated and coordinated is important (Cox and Townsend, 1997). According to a report by Constructing Excellence (2004: 4), "construction is a collaborative activity—only by pulling the knowledge and experience of many people can buildings meet the needs of today, let alone tomorrow. But simply bringing people together does not necessarily ensure they will function effectively as a team." Thus, construction participants or supply chains (e.g., clients, architects, engineers, quantity surveyors, contractors, and suppliers) are required to work effectively as a team to deliver a project successfully.

The level of integration in the construction supply chain stays is one of the major problems of the construction industry (Akintoye et al., 2000). The key barriers to integration originate from the historical fragmentation of project delivery systems and the adversarial culture of construction project relationships (Dainty et al., 2001). A fragmented and largely subcontracted workforce has increased the complexity of the construction supply chain and disabled the process of integration (Biscoe and Dainty, 2005). The relationships between the contractors and subcontractors, moreover, are heavily based on price and competitive bidding (Dainty et al., 2001). In this context, the downstream of construction does not have efficient and longstanding supplier-contractor relationships (Akintoye et al., 2000). A case study in small and medium enterprises for the construction industry (Dainty et al. 2001) revealed that there is growing interest in the integration of the upstream of construction supply chain (clients, consultants, and contractors); however, there is a lack of interest and development in the downstream (contractors and sub-contractors).

In relation to quality and productivity problems, more inter-organisational relationships and communications would reduce the inefficiencies that cause additional costs and delays in delivery arising from design changes and rework. Although inter-organisational relationships are not new, pursuing greater levels of integration in the industry will help address the challenges brought about by fragmentation. Therefore, it is perhaps not surprising that most construction projects will evidence some forms of relational contracting approaches and that the use of RC in the construction industry has grown worldwide (Colledge, 2005). The core values of the relationship rely upon commitment, trust, respect, communication, collaboration, innovation, fairness, and enthusiasm. KM through knowledge sharing and knowledge transfer is important for effective communication and collaboration. However, several barriers for such integration approaches have been identified. These include lack of a learning culture, lack of a knowledge-sharing culture, low levels of commitment amongst partners, lack of competency, focusing on the project instead of the process, and the adversarial culture of the industry (Maqsood, et al., 2003), which may hinder knowledge flow within the construction processes. Such barriers can be minimised through the effective implementation of KM, which can facilitate collaboration through tools and processes that help ensure the flow of the knowledge utilised by the project team.

### 3.2. KM and its relationship to team integration

KM has been defined in many different ways. Wiig (1997) suggests that KM deals with the management of knowledge-related activities, which include creating, organising, sharing, and using knowledge to create value for an organisation. A more formal definition of KM, given by the American Productivity and Quality Center, is the strategies and processes of identifying, capturing, and leveraging knowledge (Carrillo and Chinowsky, 2006). Von Krogh (1998) refers to KM as the process of identifying and leveraging the collective knowledge in an organisation to help it compete. However, the most common description of KM is as a business practice that emphasises the creation, dispersion, and use of knowledge (Davenport and Prusak, 1998; Alavi And Leidner, 2000). In this case, the purpose of KM is to enable the organisation to gain access to the knowledge held within the individuals of the firm.

Davenport and Prusak (1998: 5) define knowledge as “a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of knower. In organisations, it often becomes imbedded not only in documents or repositories but also in organisational routines, processes, practices, and norms.”

There are several dimensions of organisational knowledge, including individual and group knowledge, and internal and external knowledge (Al-Ghassani et al., 2002). However, one of the most practical distinctions is that between tacit and explicit knowledge (Nonaka and Takeuchi, 1995). Tacit knowledge is stored in the heads of individuals and is difficult to communicate externally or to share. On the other hand, explicit knowledge is captured or stored in an organisation’s manuals, procedures, information systems, and is easily communicated or shared with other people or parts of an organisation (Nonaka, 1991; Roberts, 2000; Robinson et al., 2005).

Within the KM literature, there are at least two classifications of knowledge processes. Nonaka and Takeuchi (1995) define one in which there are four processes: internalisation, externalisation, combination, and socialisation. The socialisation mode refers to the conversion of tacit knowledge into new tacit knowledge through social interactions and shared experiences amongst organisational members (e.g., apprenticeship). The combination mode refers to the creation of new explicit knowledge by merging, categorising, reclassifying, and synthesising existing explicit knowledge (e.g., literature survey reports). Internalisation is a process in which an individual internalises explicit knowledge to create tacit knowledge. Finally, externalisation is a process in which a person turns their tacit knowledge into explicit knowledge through documentation, verbalisation, etc. This classification focuses on different processes in which knowledge is created and transferred throughout an organisation.

Another classification focuses on the lifecycle of knowledge within an organisation (Alavi and Leidner, 2000; Davenport and Prusak, 1998; Teece, 1998). It includes knowledge generation (creation and knowledge acquisition); knowledge codification; (storing); knowledge transfer (sharing); and knowledge application. Knowledge generation involves the discovery and resolution of opportunities or problems; the creation of innovations (Gray and Chan, 2000; Matusik and Hill, 1998); and knowledge acquisition, which is the acquiring and integrating of knowledge from external sources (Davenport and Prusak, 1998). Knowledge codification is the translation of knowledge into text, drawings, etc. for storage in a repository. Knowledge transfer refers to the sharing of knowledge amongst individuals within an organisation. Finally, knowledge application is the use of knowledge to gain a competitive advantage (Alavi and Leidner, 2000).

KM is not entirely new to the construction industry. In some form, construction organisations have always managed their knowledge and relied on the expertise of key members of staff (Carrillo, 2004). KM has always been a challenge to the construction industry, which is predominantly project-based (Kamara et al., 2000). Information overload, lack of time to share knowledge, not using technology to share knowledge effectively, and difficulty capturing tacit knowledge are a few challenges in implementing KM (Carrillo et al., 2004). Other identified barriers include lack of management support, employee resistance to sharing knowledge, poor ICT infrastructure, lack of dedicated resources, poor organisational culture, poorly articulated strategy, and difficulty in evaluating benefits (Robinson et al., 2005; Dainty et al., 2005).

In the construction industry, KM is required at the inter-organisational level (i.e., within projects and across temporary and multidisciplinary project organisations), and at the intra-organisational level (i.e., within individual firms) (Kamara et al., 2002). Because it is aimed at improving integration, the present research will examine how KM is being implemented at the inter-organisational level—i.e., in a collaborative project or relational contracting (RC) setting. Implementing KM remains a challenging task for organisations. As Drucker (1993) asserts, one of the most important challenges facing organisations is to build systematic practises for managing knowledge. Therefore, to ensure the success of their KM activities, it is appropriate that a sound implementation framework be developed to guide organisations before the actual implementation takes place. The next section discusses the concept of RC in the construction industry in order to provide a better understanding of how this framework can be utilised to implement KM for integration improvement.

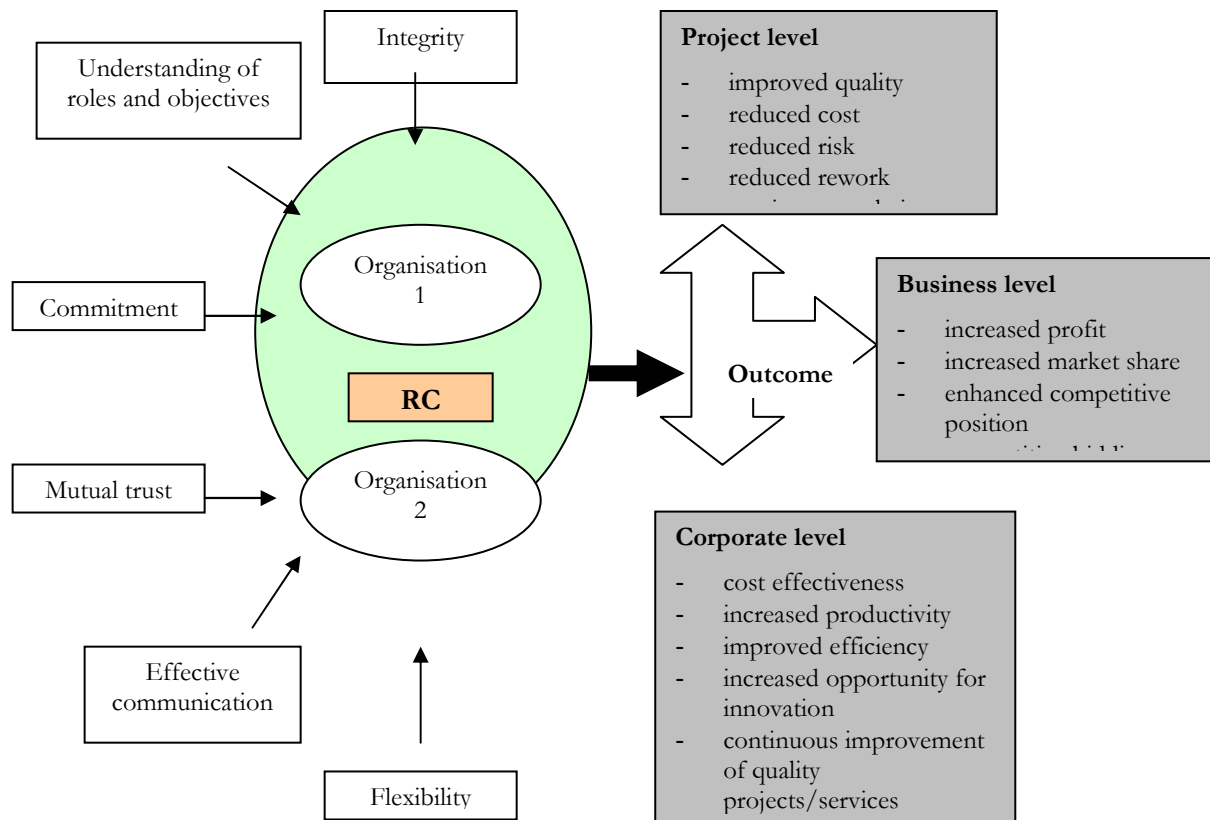
### *3.3. The concept of relational contracting (RC)*

Theories that emphasise the benefits of close, long-term relationships amongst different organisations are receiving increasing attention throughout the academic literature. One of these theories is RC. Many other terms have been used to describe relationship phenomena, such as relationship quality, cooperative relationships, partnering, strategic alliances, supply chain management, and team-working.

RC (also known as relational contract theory) has been defined as “the relations among parties to the process of projecting exchange into the future” (Macneil, 1980: 4; see also Faisal et al., 2005). RC theory is a socio-legal philosophy of contracting (Macneil and Capbell, 2001), which views the need for the enforcement of formal contracts or agreements as less import than the need to maintain relationships in the interest of future cooperation (Arrighetti et al., 1997). RC is an appropriate way to provide the necessary flexibility in contractual relationships and overcome transactional barriers to teambuilding (Rahman and Kumaraswamy, 2004a; 2004b; 2002). Colledge (2005: 31) defines RC as “a transaction or contracting mechanism that seeks to give explicit recognition to the commercial “relationship between the parties to the contract. In essence, the terms of the contract assume less importance than the relationship itself, with mechanisms for delivery that focus on trust and partnership.”

CRC CI (2002) suggested a working definition of RC that consists of the following characteristics:

- based on recognition and striving for mutual benefits
- win-win scenarios
- cooperative relationships
- underpins various approaches (partnering, alliances, joint ventures, etc.)
- better risk-sharing mechanisms
- long-term relationships develop and change over time
- involves substantial relationships amongst the parties



**Figure 1.** Relational contracting: success factors and outcomes.

Several researchers (Badger and Mulligan, 1995; Cheng et al., 2004) have highlighted the positive outcomes (see Figure 1) of building good relationships. In relational contracting, such relationships can occur at three different levels:

1. *Project level*, where the benefits may include improved quality, reduced cost, reduced risk, reduced rework, and on-time completion.
2. *Business level*, where the benefits may include increased profits, increased market share, enhanced competitive position, and competitive bidding.
3. *Corporate level*, where the benefits may include cost effectiveness, increased labour productivity, improved efficiency, increased opportunity for innovation, increased cultural responsiveness, and continuous improvement of quality products and services.

At a project level, RC can improve working relationships amongst all project stakeholders in the construction industry; facilitate efficient and effective construction; enhance financial returns; minimise the incidence of conflict; and facilitate conflict resolution (Colledge, 2005).

Some key issues that have been identified in the literature include trust (Kadefors, 2004; Swan et al., 2002, McDermott et al., 2005); and commitment, mutuality, openness, flexibility, long-term perspectives, teamwork, and honesty (Black et al., 2000; Cheng et al., 2004; Wood and Ellis, 2005). Kwawu and Hughes (2007) also suggest that the success of the relationship is basically dependent on mutual trust, commitment, and cooperation in both performance and further planning.

The need for relational approaches or inter-organisational relationships in construction might be due to the nature of the industry itself, which is often highly specialised, involving multiple participants and complex projects that require extended periods of time for commencement and completion (Colledge, 2005). Dissanayaka and Kumaraswamy (1999) note that given the nature of the construction industry, conflicts between diverse participants need to be minimised through better relationships and cooperative teamwork. The shift towards more RC relationships has been evident in the increased use of project partnering agreements, together with the development of construction process relational tools such as project team goals, meetings, and reviews (Colledge, 2005). There have been calls for radical cultural changes to construction procurement in the UK (Murray and Langford, 2003). RC has been proposed as a

substitute for traditional or formal contracting, with the aim of overcoming inadequate organisational systems, adversarial contractual relationships, and mistrust in the UK construction industry (Latham, 1994; Egan, 1998).

Previous research on inter-organisational relationships in the construction industry has been conducted from various perspectives. Most of these studies have concentrated on the relationships between the principal participants in the industry—namely the client, main contractor, subcontractor, supplier, and consultant (Faisal et al., 2006). As a result, many terminologies have been developed and used in different contexts. For example, the term “partnering” has been widely used to refer to “alliances” within the supply chain. Such alliances are usually informal (rather than contractual) relationships (Bresnen and Marshall, 2000; Cheng et al., 2004). The importance of embedding the relationship within a contract has been recognised, resulting in the emergence of new forms of relationships, such as public-private partnerships (Ahadzi and Bowles, 2004; Parker and Hartley, 2003). In addition, partnering has been defined in many ways. It can be viewed as an individual project mechanism or as a long-term strategy (Faisal et al., 2006). Alliancing is normally assumed to be a long-term business strategy, linking together client, contractor, and supply chain (Rowlinson and Cheung, 2004). Gulati (1998) defined an alliance as any voluntarily initiated cooperative agreement between firms that involves exchange, sharing, or co-development; it can include contributions by partners of various resources. RC provides a more efficient and more effective contracting mechanism for certain types of transactions, particularly when close collaboration amongst parties is required in order to realise complex construction projects or long-term development programmes (Colledge, 2005).

A common approach of RC is partnering or collaborative contracting. This tool encourages the creation, maintenance, and utilisation of a knowledge pool across organisations. However, how to facilitate KM within the collaborative work culture is an under-researched area.

#### *3.4. The link between RC and KM*

RC and KM are two significant management concepts. Both concepts focus on communication, trust, knowledge sharing, loyalty, and commitment (Rowley, 2004). Repositories of data or information can be used as a platform for processes associated with both relationships and knowledge. Both RC and KM recognise the value to be created through appropriate synergies of technology, people, and process. KM principles and techniques play an important part in the success of RC. Some of its basic principles include a systematic approach for capturing, codifying, and sharing information and knowledge; a focus on building social capital to enable collaboration amongst people and communities; an emphasis on learning and training; and an emphasis on leveraging knowledge and expertise in work practice (Khamseh and Jolly, 2008).

KM is primarily concerned with capturing, codifying, transferring, and sharing both tacit and explicit knowledge (Nonaka, 1994)). The challenge of KM is to convert tacit knowledge into explicit knowledge through the balanced use of technology and soft human-related factors, such as leadership, vision, strategy, reward systems, and culture. Briefly, insufficient knowledge flow amongst construction participants, inefficient communication channels, inability to share information, and lack of a knowledge-sharing culture and common goals can be regarded as the barriers to effectively implementing KM in a collaborative environment (Briscoe et al., 2001). Therefore, it is important to study the critical factors that influence the success of KM in organisations, so that they can be extended to improve integration within the RC setting. The similarities between KM and RC could be used as a basis for developing a new framework that employs CSFs for KM as tools for improving integration in the construction industry.

#### *3.5. The concept of critical success factors (CSFs)*

Organisations have used critical success factors (CSFs) to identify the key organisational capacities necessary for success and to focus their efforts on building these capabilities.

Several definitions of CSF are provided in the literature. Drawing on ideas from Daniel (1961) and Anthony et al. (1972), Rockart (1979: 85) provides one of the most frequently cited definitions: CSFs, he notes, refer to “the limited number of areas in which results, if they are satisfactory, will ensure successful competitive performance for the organization. They are the few key areas where things must go right for the business to flourish. If the results in these



areas are not adequate, the organization's efforts for the period will be less than desired." Consequently, Rockart (1979) stresses that companies should constantly and carefully manage these particular areas of activity. In a similar vein, Bruno and Leidecker (1984: 24) define CSFs as "those characteristics, conditions or variables that, when properly sustained, maintained, or managed, can have a significant impact on the success of a firm competing in a particular industry." Pinto and Slevin (1987: 22) regard CSFs as "factors which, if addressed, significantly improve *project implementation chances*." However, according to Esteves (2004), the latter two definitions fail to address the comprehensive concept proposed by Rockart (1979), which emphasises the importance of creating an ideal match between environmental conditions and the business characteristics of a particular company.

CSFs can be events, conditions, circumstances, or activities. Specifically, they are "the limited number of areas in which results, if they are satisfactory, will ensure successful competitive performance of an operation" (Jenster, 1987: 103). The identification of these factors provides a vehicle for the design of an effective system of performance measurement and control. The CSF method is considered a top-down approach for planning (Byers and Blume 1994). CSFs are areas that should receive constant and careful attention and monitoring. These factors are also useful in prioritising critical concerns for a project. In addition, the CSF method generates user acceptance amongst senior management, and works well at the policy, operational, and strategic levels of information resource planning (Dobbins and Donnelly, 1998).

A considerable amount of literature has been published on CSFs for KM. Since the late 1990s, many researchers have attempted to develop a comprehensive list of factors. However, the lists differ because of the multidisciplinary nature of KM, and because of the different backgrounds and interests of KM researchers. Our literature review that found that CSFs for KM can be broadly categorised into three aspects: people, process, and technology, as shown in Figure 2:



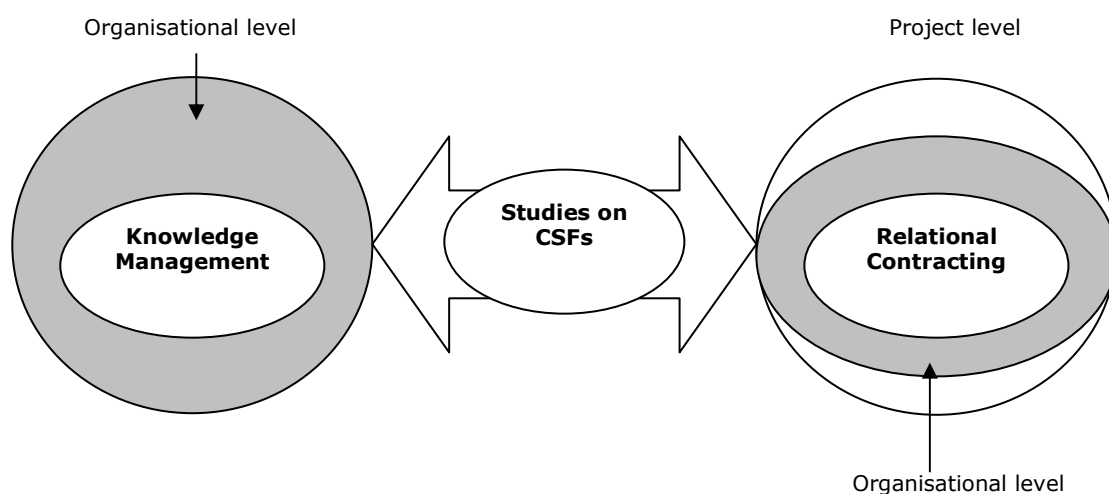
**Figure 2.** Summary of CSFs for KM.

Moreover, CSFs vary according to the scope and context of the relevant research. It was found that the CSFs identified focus mainly on the organisational level.

Only several studies have been done to identify the CSFs for KM in the construction industry. Bishop et al. (2008) identifies a set of people-oriented CSFs that ensure the effectiveness of the KM initiative before, during, and after their implementation in the context of UK-based construction organisations. Another study by Lin and Lin (2006) identifies several CSFs for implementing KM in construction projects from the perspective of the oil and gas sector in Taiwan. In general, therefore, it seems that more studies need to be conducted to examine the CSFs for KM within construction, which is considered to be a project-based industry.

RC is a very broad concept for an inter-organisational relationship. Our literature review of studies on CSFs was thus based on various approaches, such as partnering, strategic alliances, public-private partnerships, and joint-ventures. Common CSFs for RC-based approaches include mutual trust (Kwawu and Hughes, 2007; Kadefors, 2004; Swan et al., 2002; McDermott et al., 2005); long-term commitment, support from top management, effective communication, teamwork, motivation, and coordination (Wood and Ellies, 2005; Cheng et al., 2004; Dainty et al., 2001; Miller et al., 2001; Bresnen and Marshall, 2000).

Previous studies of CSFs in the construction industry suggest that instead of examining organisational management practices, the identification of CSFs has been focused on the project level. The adoption of a bottom-up approach in the construction industry (versus a top-down approach in other industries) has limited the development of construction organisations as business practice leaders. Therefore, it is important to examine how research on CSFs for KM, which has been conducted at the organisational level, can be further extended to the project level within the RC project setting. This argument about the relationship between the two findings is illustrated in Figure 3.



**Figure 3.** The context of CSFs studies.

#### 4.0. Research methodology

A comprehensive review of the issues relating to the fragmentation and lack of integration in the construction industry was conducted in order to highlight the need for the construction industry to use a collaborative approach such as relational contracting (RC) to improve integration within construction project teams. The key aims of integration include improving trust and long-term relationships amongst project participants, encouraging knowledge sharing and knowledge transfer, and improving communication. A review of literature on the concepts of Knowledge Management (KM) was conducted to gain insight into how the KM can facilitate integration during the construction process through the use of its critical success factors (CSFs). This literature review will form the basis of a conceptual framework to be related to case study projects in ongoing doctoral research into how improved integration can be established throughout the construction supply chain.

#### 5.0. Conclusion and the way forward

Organisations need to value and manage both knowledge and relationships in order to improve integration amongst construction players. It is important to understand the interaction between the complementary paradigms of KM and RC. Clearly, KM initiatives and the great potential of using RC to assist construction industry clients and practitioners in overcoming the fragmentary and adversarial nature of contracting are not being fully realised in practice. In conclusion, CSFs can be a useful tool to help identify important and critical issues of both RC and KM, and thereby improve integration within the construction supply chain.

The next stage of the research is to develop the conceptual framework in order to show the relationship between the concepts of CSFs for KM and the roles of KM in improving integration within the RC project settings. Then, a pilot study will be carried out to verify the CSFs for KM identified in literature. This will be followed by the main data collection using the case study method on RC projects in Malaysia.

## References

- Adnan, H. and Morledge, R. (2003). 'Critical success factors in Malaysian construction joint venture projects,' *Proceedings of the RICS Foundation Construction and Building Research Conference (COBRA)*.
- Ahadzi, M. and Bowles, G. (2004) 'Public-private partnerships and contract negotiations: an empirical study,' *Construction Management and Economics* 22 (9), pp. 967–978.
- Alavi, M. And Leidner, D.E. (2001), 'Review: knowledge management and knowledge management systems: conceptual foundations and research issues', *MIS Quarterly* 25 (1), pp. 107–36.
- Amaratunga, D. et al. (2002) 'Quantitative and qualitative research in the built environment: application of "mixed" research approach,' *Work Study*, 51 (1), pp. 17–31.
- Amaratunga, D. and Baldry, D. (2001) 'Case study methodology as a means of theory building: performance measurement in facilities management organisations,' *Work Study*, 50 (3), pp. 95–104.
- Bakri, A. and Saidan-Khaderi, S. (2007) 'Knowledge management practices in quantity surveying firms in Malaysia,' *Proceedings of Quantity Surveying International Convention (QSIC)*, Kuala Lumpur, Malaysia.
- Bresnen, M. and Marshall, N. (2000) 'Building partnerships: case studies of client-contractor collaboration in the UK construction industry,' *Construction Management and Economics* 18 (7), pp. 819–832.
- Bresnen, M., Edelman, L., Swan, J., Laurent, S., Scarborough, H., and Newell, S. (2002) 'Cross-sector research on knowledge management practices for project-based learning,' *Proceedings of the European Academy of Management 2<sup>nd</sup> Annual Conference*, Sweden.
- Bruno, A. and Leidecker, J. (1984) Identifying and using critical success factors. *Long Range Planning*, 17 (1), pp. 23-32
- Cavana, R.Y. et al. (2001) *Applied Business Research: Qualitative and Quantitative Methods*: Australia: John Wiley & Sons.
- Chan, A.P.C., Chen, D.W.M., Chiang, B.S.T., Chan, E.H.W., and Ho, K.S.K. (2004). 'Exploring critical success factors for partnering in construction projects. *Journal of Construction, Engineering and Management*, ASCE 130 (2), pp. 188–198.
- Cheng, E.W.L., Li, H., Love, P.E.D., and Irani, Z. (2004) 'Strategic alliances: a model for establishing long-term commitment to inter-organisational relations in construction,' *Building and Environment* 39 (4), pp. 459–468.
- Colledge, B. (2005). 'Relational contracting: creating value beyond the project,' *Lean Construction Journal* 2 (1), pp. 30–45.
- Constructing Excellence (2004) Demonstrating Excellence – An Evolution of the Programme of Demonstrations, *Constructing Excellence*.
- Creswell, J.W. (2003) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 2<sup>nd</sup> Edition. London: Sage Publications.
- CRC CI (2002) 'A review of the concepts and definitions of the various forms of Relational Contracting (2002-022-a). Cooperative Research Centre (CRC) Construction Innovation, Brisbane, Australia.
- Dainty, A.R.J., Briscoe, G.H. and Millet, S.J. (2001) 'New perspectives on construction supply chain integration,' *Supply Chain Management* 6 (4), pp. 163–173.
- Davenport, T.H. and Prusak, L. (1998), *Working Knowledge: Managing What Your Organisation Knows*. Boston, MA: Harvard Business School Press.

- Davenport, T.H., De Long, D.W., and Beers, M.C. (1998) 'Successful knowledge management projects,' *Sloan Management Review* 39 (2), pp. 43–57.
- Dissanayaka, S.M. and Kumaraswamy, M.M. (1999) 'Reconstructing procurement systems and team relationships,' *International Journal of Computer Integrated Design and Construction* 1 (2), pp. 10–19.
- Egan, J. (1998) *Rethinking Construction: The report of the Construction Task Force to the Deputy Prime Minister, John Prescott, on the Scope for Improving the Quality and Efficiency of UK Construction*. London: Department of the Environment, Transport, and the Regions Construction Task Force.
- Easterby-Smith, M. et al. (2003) *Management Research: An Introduction, 2<sup>nd</sup> Edition*. London: Sage Publications
- Faisal, N., Dainty, A.R.J., and Price, A.D.F. (2005) 'The concept of "relational contracting" as a tool for understanding inter-organizational relationships in construction,' in Khosrowshahi, F., ed., *Proceedings of the 21<sup>st</sup> Annual Association of Researchers in Construction Management Conference 2*, SOAS, London, pp. 1075–84.
- Faisal, N., Dainty, A.R.J., and Price, A.D.F. (2006) 'Perceptions of construction organisations on developing successful inter-organisational relationships,' in Boyd, D., ed., *Proceedings of the 22<sup>nd</sup> Annual Association of Researchers in Construction Management Conference 1*, Birmingham, UK, pp. 471–479.
- Fong, S.W.P. (2005) 'Building a knowledge-sharing culture in construction project teams, in Anumba, C.J, Egbu, C., and Carrillo, P., eds., *Knowledge Management in Construction*. Oxford: Blackwell Publishing.
- Gulati, R. (1998) 'Alliances and networks,' *Strategic Management Journal* 19 (4), pp. 293–317.
- Hasanali, F. (2002) 'Critical success factors of knowledge management'; available at: [www.kmadvantage.com/docs/km\\_articles/CriticalSuccess\\_Factors\\_of\\_KM.pdf](http://www.kmadvantage.com/docs/km_articles/CriticalSuccess_Factors_of_KM.pdf).
- Ingirige, B. and Sexton, M. (2006) 'Alliances in construction: investigating initiatives and barriers for long-term collaboration,' *Engineering, Construction and Architectural Management* 13 (5), pp. 521–535.
- Jefferies, M., Gameson, R. and Rowlinson, S. (2002) 'Critical success factors of the BOOT procurement system: reflection from the Stadium Australia case study,' *Engineering, Construction and Architectural Management* 9 (4), pp. 352–61.
- Jenster, P.V. (1987) Using critical success factors in planning. *Long Range Planning*, 40, pp 102-109
- Kaderfors, A. (2004) 'Trust in project relationships: inside the black box,' *International Journal of Project Management* 22 (3), pp. 175–182.
- Kamara, J.M., Chimay, A.J., and Carrillo, P.M. (2002) 'A CLEVER approach to selecting a knowledge management strategy,' *International Journal of Project Management* 20 (3) pp. 205–211.
- Khamseh, H.M. and Jolly, D.R. (2008) 'Knowledge transfer in alliances: determinant factors,' *Journal of Knowledge Management* 12 (1), pp. 37–50.
- Kogut, B. (1988) 'Joint ventures: theoretical and empirical perspectives,' *Strategic Management Journal* 9 (4), pp. 319–32.
- Kwawu, W. and Hughes, W. (2007) 'Are relational and formal contracts implemented as substitutes in the UK construction industry?' in Boyd, D., ed., *Proceedings of the 23<sup>rd</sup> Annual Association of Researchers in Construction Management Conference 1*, Belfast, UK, pp. 97–106.
- Latham, M. (1994) *Constructing the Team: Final Report of the Government/Industry Review of Procurement and Contractual Arrangements in the UK Construction Industry*. London: HMSO.
- Li, B., Akintoye, A., Edwards, P.J., and Hardcastle, C. (2005) 'Critical success factors for PPP/PFI projects in the UK construction industry,' *Construction Management and Economics* 23 (5), pp. 459–471.

- Liebowitz, J. (1999) 'Key ingredients to the success of an organisation's knowledge management strategy,' *Knowledge and Process Management* 6 (1), pp. 37–40.
- Macneil, I.R. (1974) 'The many futures of contracts,' *Southern California Law Review* 47, pp. 691–816.
- Maqsood, T. and Walker, D. (2007) 'Extending the "knowledge advantage": creating learning chains,' *The Learning Organisation* 14 (2), pp. 123–141.
- McDermott, P., Khalfan, M.M.A., and Swan, W. (2005) 'Trust in construction projects,' *Journal of Financial Management of Property and Construction* 10 (1), pp. 19–31.
- Mohamed, O., Othman, M., Mahanim, H., Wan Omar, W.M.S., and Mohd Nawawi, M.N. (2007) 'Knowledge sharing: problems and barriers to implementation,' *Journal of the Surveyors, Malaysia* 40 (3), pp. 27–31.
- Morledge, R. and Owen, K. (1998) 'Critical success factors in PFI projects,' *Proceedings of the 14<sup>th</sup> Annual Association of Researchers in Construction Management Conference*, The University of Reading, UK, pp. 565–574.
- Morledge, R. and Owen, K. (1999) 'Developing a methodological approach to the identification of factors critical to success in privatised infrastructure projects in the UK,' in Ogunlana, S., ed., *Profitable Partnering in Construction Procurement*. New York: Routledge.
- Narayanan, R., Richardson, S., and Salleh, A.L. (2003) 'Knowledge management: a Malaysian study,' *Malaysian Management Review* 38 (1), pp. 79–87.
- Nonaka, I. (1994) 'A dynamic theory of organisational knowledge creation,' *Organisation Science* 5 (1), pp. 14–37.
- Nonaka, I. and Takeuchi, H. (1995) *The Knowledge Creating Company: How Japanese Companies Create Dynamics of Innovation*. Oxford: Oxford University Press.
- Rahman, M.M. and Kumaraswamy, M.M. (2002) 'Joint risk management through transactionally efficient relational contracting,' *Construction Management and Economics* 20 (1), pp. 45–54.
- Rahman, M.M., Kumaraswamy, M.M., and Ng, S.T. (2003) 'Reengineering construction project teams,' in Molenaar, K.R. and Chinowsky, P.S., eds. *Proceedings of the American Society of Civil Engineers Construction Research Congress*, Hawaii.
- Remenyi, D. et al. (2003) *Doing Research in Business and Management: An Introduction to Process and Methods*. London: Sage Publications.
- Rockart, J. (1982) 'The changing role of the information systems executive: a critical success factors perspective,' *Sloan Management Review* 23 (1), pp. 3–13.
- Rowley, J. (2004). 'Partnering paradigms? knowledge management and relationship marketing,' *Journal of Industrial Management & Data Systems* 104 (2), pp. 149–157
- Saunders, M. et al. (2007) *Research Methods for Business Students, 4<sup>th</sup> Edition*. Edinburgh Gate: Pearson Education Ltd.
- Sexton, M (2003) 'A supple approach to exposing and challenging assumptions and path dependencies in research,' *3<sup>rd</sup> International Postgraduate Research Conference*, Lisbon.
- Siemieniuch, C.E. and Sinclair, M.A. (1999) 'Organisational aspects of knowledge life-cycle management in manufacturing,' *International Journal of Human-Computer Studies* 51 (3), pp. 517–547.
- Strategic Forum for Construction (2003). *The Integration Toolkit Guide: Integrated Project Team*. London: Strategic Forum for Construction.
- Sundaraj, G. (2007) 'The way forward: construction industry master plan 2006–2015,' *Master Builders*, 1<sup>st</sup> Quarter 2007; available at:  
[http://www.mbam.org.my/mbam/images/MBJ1Q07/@WayForward\(48\).pdf](http://www.mbam.org.my/mbam/images/MBJ1Q07/@WayForward(48).pdf).
- Swan, W., McDermott, P., Wood, G., Abott, C., and Thomas, A. (2002) *Trust in Construction: Achieving Cultural Change*. Salford: Centre of Construction Innovation.
- Tiong, R.L.K. et al. (1992) 'Critical success factors in winning BOT contracts,' *Journal of Construction Engineering and Management* 118 (2), pp. 217–228.

- Welman, C., et al. (2005) *Research Methodology. 3<sup>rd</sup> Edition*: Oxford University Press, Southern Africa
- Wood, G. D. and Ellis, R. C. T. (2005) 'Main contractor experiences of partnering relationships on UK construction projects,' *Construction Management and Economics* 23 (3), pp. 317–325.
- Wong, K.Y. (2005) 'Critical success factors for implementing knowledge management in small and medium enterprises,' *Industrial Management & Data Systems* 105 (3), pp. 261–279.
- Wong, K.Y. and Aspinwall, E. (2005) 'An empirical study of the important factors for knowledge-management adoption in the SME sector,' *Journal of Knowledge Management*, 9 (3), pp. 64–82.
- Yin, R.K. (2003) *Case Study Research: Design and Methods, 3<sup>rd</sup> Edition*. London: Sage Publications.

# The Effect of Organisational Structure on Communication Behaviour in Libyan Public Organisations

**Salh Bezweek and Charles Egbu**, *School of the Built and Human Environment, University of Salford*

## Abstract

This paper presents a critical review of the literature pertaining to the effect of organisation structure on communication behaviour in public organisations in Libya. It identifies deficiencies in this area of study and draws lessons for future research. The focus is primarily on the effect of three structure dimensions—centralisation, formalisation, and complexity—on communication behaviour in organisations. Key issues associated with the definitions of structure, communication behaviour, and organisational culture are considered. In addition, a critical review and synthesis of related studies on the effect of structure dimensions on the communication behaviour of individuals in organisations is provided. In conclusion, the paper argues that the effect of organisation structure on communication behaviour in organisational settings is complex. It also emphasises the “tall” character of organisational structures in Libya, which are marked by complex hierarchal, bureaucratic, and central decision-making processes. The implications of these issues for research strategy and design are noted, along with recommendations for future research.

**Keywords:** Communication behaviour, Libyan national culture, organisational structure, public organisations

## 1. Introduction

The theoretical and intuitional field of communication studies has developed over the past three decades, during which time various problems, issues, and debates have emerged in the literature. This review considers the influence of organisational structure on the communication behaviour of individuals in Libyan organisations. Recent structure studies are examined in order to investigate the impact of organisational structure on communication amongst staff in Libyan public organisations and to provide suggestions for future research on this issue. The paper also analyses the extent to which formality, centrality, complexity, and stratification affect the communication behaviour of individuals in Libyan organisations.

## 2. Methodology

This paper draws on an ongoing doctoral study of communication behaviour and quality in Libyan public organisations. It is based on a thorough critical review of academic journals, conference proceedings, textbooks, and websites related to communication behaviour, organisational structure, and power influences in Libyan public organisations. The paper’s aim is to study the effect of organisational structure on communication behaviour in these organisations.

## 3. Communication

“Communication can be defined as the process by which information is exchanged and understood by two or more people, usually with the intent to motivate or influence behaviour” (Daft, 1997: 560). Communication in an organisation can be undertaken in a number of different directions—i.e., downwards, upwards, etc. (Udall and Udall, 1979; Daft, 1997). For example, at the senior level managers spend 80% of their time communicating (Daft, 1997: 583). There is often horizontal communication at meetings, but problems can arise further down the chain at the lower levels. How communication is managed here is an important question to address, because at this level it is more likely that the “grapevine” will play a part in communication (Udall and Udall, 1979). Many experts regard downward communication—

which refers to those messages that pass from the top down—as the most familiar. Daft (1997: 569) observes “that different employees are responsive to different kinds of communication.” He also notes that the main problem with downward communication is the distortion or loss of the message (up to 25%) as it passes through the chain.

In contrast, in upward communication, information flows from the lower to higher levels of an organisation (Daft, 1997: 570). Daft identifies a number of examples of such communication, including suggestion boxes, employee surveys, open-door policies, and face-to-face conversations. Many of these communication avenues are currently utilised by local councils in Northern Ireland. However, methods may vary from council to council and barriers do exist, such as the unwillingness of managers to listen or employees not trusting managers sufficiently to put forward suggestions for improvements. According to Daft (1997: 570), upward communication entails five key elements: problems and exceptions; suggestions for improvement; performance reports; grievances and disputes; financial and accounting information.

A third type of internal communication—identified by Jefkins (1998)—is one that takes place amongst employees in a “sideways” direction.

#### **4. Organisational communication**

Organisational communication consists of various ways of sending and receiving messages; it affects the formal social units in which individuals work towards common goals. It is generally identified with written media (correspondence, house publications, bulletin board information); hardware (telephone systems, dictation equipment, computer units); and speech activities (interviewing, directing, conferring). However, it should be stressed that organisational communication includes all behaviour-modifying stimuli, both verbal and nonverbal. It includes gestures and facial expressions; words spoken in conversations and meetings; symbols and colours; and words written in reports. Therefore, organisational communication is concerned with most of the specialised disciplines related to human communication—i.e., interpersonal communication, small-group communication, information systems, mass communication, and nonverbal communication.

There are many definitions of organisational behaviour in the literature. Gibson et al. (2000: 5) describe organisational behaviour as follows: “The field of study that draws on theory, methods and principles from various disciplines to learn about individual perspectives, values, learning capacities and actions while working in groups and within the total organisation; analysing the external environments effect on the organisation and it’s human resources, mission, objectives and strategies.” Daft and Noe (2001: 4) define organisational behaviour as: “The actions and interactions of individuals and groups in organisations.” Communication is widely studied as a means of transmitting ideas as part of culture (Jandt, 2004). Therefore, organisational communication can be defined as a system in terms of purpose, operational procedures, and structure. The purpose of organisational communication is to facilitate the achievement of organisational goals.

Communication can be divided into three types: verbal (use of words with specific meaning); paraverbal (tone of the voice); and nonverbal. Nonverbal communication can be defined as communication without words, while verbal communication is defined as communication with words (Remland, 2004). In addition, success in communication involves various components, such as encoding, message, channel, receiver, decoding, and receiver responses.

##### *4.1. Verbal communication*

Communication and culture are strongly related to each other. Thus, the study of communication in Western culture has a recorded history of some 2,500 years. Aristotle’s *Rhetoric* and *Poetics* describe the process of communication as involving a speaker, the speech act, an audience, and a purpose (Jandt, 2004). The basis of communication is the interaction between people. Verbal communication is one way for people to communicate face-to-face. Some of the key components of verbal communication are sound, words, speaking, and language. The main element of communication is language. All languages are similar in linguistic structure and express different cultural aspects (Fisher, 1978). According to Wary



and Grace (2007), cultural diversity plays an important role in misunderstanding in exoteric communication. The ways in which communication is approached and studied will vary in accord with the characteristics of the particular culture under consideration. The understanding of the hearer is the responsibility of speaker, who must take further steps to clarify his or her message. It is therefore important to understand the cultural background of the listener and to clarify the meaning of his or her message in order to facilitate communication (Jandt, 2004).

#### *4.2. Nonverbal communication*

Nonverbal communication is a very important carrier of information. It includes facial expressions, eye contact, tone of voice, emotions, attitudes, feelings expressed in different gestures, and body language (Hargie and Dickson, 2004). Its role is highly significant and essential in social interactions. One of the most important nonverbal signals in social interactions is eye contact. Eye gaze is defined by Argyle (1988, P 153) as "the meeting of two people's gaze or their looking into each other's eyes." Furthermore, Argyle (1988, P162) suggests that "People look more at those they like." In contrast, reduction of gaze is often a sign of disapproval, a lack of control, ignorance, or a lowered level of intimacy, depending on the context of the interaction. Facial expression is a display of affect; it shows the emotional state of an individual. An inappropriate facial expression may block effective communication. For example, in Libyan culture most people use certain facial expression to express fear, sadness, happiness, and surprise (Twatti, 2006). These facial expressions are less likely to be understood by people from different cultural backgrounds (Samover and Porter, 2004). The face-saving strategy, for instance, is a common feature of communication in Libyan organisations. This may be due to the emphasis that Libyan culture places on the reputation of an individual's name, family, and tribe. Indeed, societal reputation is a very important element for relationships in Libyan society as a whole (Twati, 2006).

### **5. Organisational structure**

Mullins (1993) and Mabey et al. (2001) describe the structure of an organisation as the pattern of the relationships between its various roles and parts. This structure serves to allocate work and responsibilities. It enables managers to plan, direct, organise, and control activities, so that the organisation's goals can be achieved.

Robbins (1994) further described the organisational structure as "the group of the functions and relationships that formally determine the functions that each unit ought to complete and the communication way among each unit." According to Robbins, the organisational structure stipulates how tasks are to be allocated, who reports to whom, and the formal coordinating mechanisms and interaction patterns that will be followed. Advancing a similar definition, Ivancevich and Matteson, (2001) observe that managers must take into account four decisions levels when designing and organisational structure: (1) division of the overall task into smaller jobs; (2) distribution of authority amongst those who are carrying out the task; (3) a basis by which individual jobs are to be grouped together; and (4) the appropriate size of the groups reporting to each superior.

#### *5.1. "Tall" versus "flat" structures*

According to Klein (2001), a tall organisational structure has many hierarchal levels. There is a long distance between the topmost manager and the bottommost manager, as well as a large number of job titles and employee career paths. The fewer subordinates a particular manager is required to guide, the taller the organisational structure. In contrast, flat structures are less bureaucratic, and their decision-making processes are decentralised and more empowering to workers (Lok and Crawford 2001). Flat structures are also less hierarchical, with a wide span of management. Flattening an organisation's structure speeds up the decision-making process, shortens lines of communication, and creates savings. However, too wide a management span may result in managers who are overextended and subordinates who receive too little guidance or control. When this happens, managers may be pressured to ignore or condone serious errors. On the other hand, if the management span is too narrow, some managers may be underutilised. Thus, the extent of the division of work, the nature of the delegation of authority, the process of departmentalisation, and the requirement of effective supervision

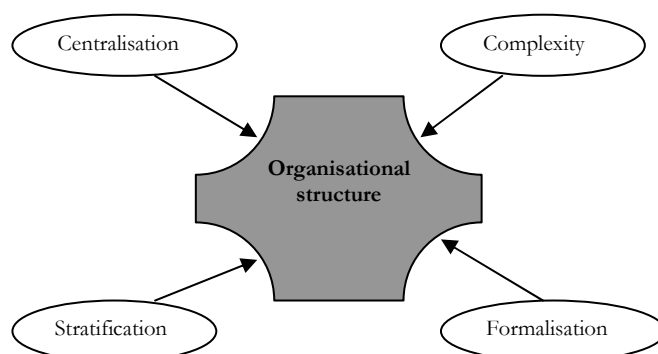
(i.e., the span of control) are all important factors to consider when designing an organisational structure.

To summarise, a tall, or hierarchal, organizational structure results from a narrow span of control, whereas a flat, or horizontal, structure is a consequence of a wide span of management. Organisations with tall structures are more inflexible and “layered.” As a result, communication within them can be difficult, as members are separated from one another within a complex and sometimes impenetrable framework (Robinson, 1990).

## 5.2. Dimensions of the organisational structure

Reimann (1973) analysed organisational structures based on research produced by the Aston group. He proposed a model with three dimensions: centralisation of authority, specialisation, and formalisation. Reimann subsequently identified a fourth dimension of lesser importance: the size of the staff.

In other models—such as those developed by Van den Ven (1976), Jackson and Morgan (1982), Fredrickson (1984), and Robbins (1994)—three additional dimensions of organisational structure are emphasised: formalisation, complexity, and centralisation. (See Figure 1.) These have been used in subsequent research undertaken from the standpoint of contingency theory.



**Figure 1.** Dimensions of the organisational structure.

### 5.2.1. Complexity

Complexity refers to the degree of differentiation within an organisation. A large variety of jobs and units creates a more complex organisational structure. Many studies have demonstrated a strong relationship between complexity and communication (Hatch, 1997). More complex organisations are made up of many horizontal and vertical divisions, and therefore have a greater need for efficient communication. Complexity is directly related to various difficulties in the transmission and reception of information, to distortion in communication, and to the loss of control of information by upper management (Galbraith, 1973). Therefore, the more complex the organisation, the more difficult communication becomes, and the more important is the need for efficient communication.

### 5.2.2. Formalisation

Formalisation refers to the level at which rules, regulations, policies, and explicit procedures govern organisational activities. Indicators of an organisation’s level of formalisation can be found in its written policies, job descriptions, procedures manuals, information systems for management, technical systems, and official lists of rules and regulations (Miller and Droge, 1986).

Formalisation tends to reduce workers’ overall discretion over their work activities, while increasing managers’ sense of control over their employees. As a result, the character of the organisation becomes impersonal, as well as lacking in flexibility and spontaneity (Zaltman,

1979). In fact, the more inflexible, formalised rules and regulations are used to govern an organisation's behaviour, the more difficult communication within it becomes.

### *5.2.3. Centralisation*

Centralisation refers to the extent to which authority and decision making are concentrated at the top. In a centralised organisation, final decisions are made exclusively by managers of the highest level, resulting in the tendency to minimise the participation of employees of an inferior category. A high level of centralisation appears to restrict channels of communication, and to inhibit an individual's capacity to generate ideas and share information and expertise with others. In decentralised organisations, decisions are made by the individuals who are closest to given situations and problems, reflecting the increased level of trust in the ability of employees to engage in decision making.

Research suggests that in decentralisation organisations promote higher levels of communication, participation, and satisfaction; however, coordination and control become more awkward (Grinyer and Ardekani, 1980). Such organisations may be more effective and adaptive because decentralisation shortens communication channels and allows for flexibility and responsiveness. On the downside, decentralisation may make it difficult to avoid chaos, inconsistency, and duplicate efforts, especially within large, complex organisations (Alder, 1999).

## **6. Libyan public organisations**

Organisations in Libya are most commonly characterised by a tall bureaucratic structure. As noted, such structures are generally inflexible, impersonal, highly standardised, and hierarchical. They are host to a large number of job titles and employee career paths; there is a long distance between the topmost and the bottommost manager, and fewer managers are available to supervise employees. These conditions that tend to impede communication within the organisation. Libyan public organisations are government run, policy driven, and controlled by general managers. Decision making is highly centralised and there is little worker empowerment: employees are enjoined to follow instructions without questions. Communication tends to be top-down, and promotion within the organisation is often based upon family ties and social relations. Leadership tends to be based on position, authority, and seniority. For example, a top employee's commitment to the organisation is strongly associated with his or her personal loyalty to the top boss.

## **7. Conclusion**

From our review of the literature, it can be concluded that Libyan organisations are characterised by tall organisational structures, which are marked by complexity, centrality, and formality. Structure is an important determinant of many facets of organisational behaviour, including communication. It is hoped that this paper will stimulate new dialogues and debates on the relationship between structural issues and communication within Libyan organisations. More empirical studies of this problem are needed, as it has not been sufficiently researched to date. In our view, future research should place particular emphasis on the potential for information technology to eliminate layers of hierarchical structure and improve organisational performance and communication in Libyan organisations.

## References

- A.H. Van den Ven (1973) 'A framework for organizational assessment,' *Academy of Administrative Science Quarterly* 18, pp. 462–476.
- Argyle, M. (1988) *Bodily Communication, 2nd Edition*. New York: Methuen
- Fisher, B.A. (1978). *Perspectives on Human Communication*. New York: Macmillan.
- Fredrickson, J.W. (1984) 'The effect of structure on the strategic decision process,' *Academy of Management Proceedings*, 11(2) pp 280-297
- Galbraith, J. (1973) *Designing Complex Organizations*. Reading, UK: Addison-Wesley.
- Grinyer, P.H. and Yasai-Ardekani, M. (1980) 'Dimensions of organizational structure: a critical replication,' *Academy of Management Journal* 23 (2), pp .405–421.
- Grinyer, P.H., Ardekani, M.Y., Al-Bazzaz, S. (1980) 'Strategy, structure, the environment and financial performance in 48 UK companies,' *Academy of Management Journal* 23 (2), pp. 193–220.
- Hargie. O. and Dickson, D. (2004) *Skilled Interpersonal Communication: Research, Theory and Practice, 4<sup>th</sup> Edition*. London: Routledge.
- Hatch, M. J. (1997) *Organization Theory: Modern, Symbolic-Interpretive and Postmodern Perspectives*. Oxford: Oxford University Press.
- Ivancevich, J. & Matteson, M. (2001). *Organizational behavior and management* (6th ed). McGraw-Hill College, Blacklick, Ohio
- Jackson, J.H. and Morgan, C.P. (1982) *Organization Theory: A Macro Perspective for Management*. Englewood Cliffs, NJ: Prentice-Hall.
- Jandt. F. (2004) *An Introduction to Intercultural Communication, 4th Edition*. London: Sage Publications.
- Jefkins, F. and Yadin, D. (1998) *Public Relations, 5<sup>th</sup> Edition*. Essex, UK: Pitman Publishing.
- Klein, E.E. (2001) 'Using Information Technology to Eliminate Layers of Bureaucracy,' *The National Public Accountant* June.
- Lok, P. and Crawford, J. (2001) 'Antecedents of organizational commitment and the mediating role of job satisfaction,' *Journal of Management Psychology* 16 (8), pp. 594–613.
- Mabey, C., Salaman, J., and Storey, J. (2001) 'Organisational structuring and restructuring," in Salaman, G., ed., *Understanding Business Organisations*. London: Routledge.
- Miller D. and Droge, C. (1986). 'Psychological and traditional determinants of structure,' *Administrative Science Quarterly* 31 pp. 539–560.
- Mullins, L.J. (1993) *Management and Organizational Behaviour, 3<sup>rd</sup> Edition*. London, Pitman Publishing.
- Adler, P.S. (1999) 'Building better bureaucracies,' *The Academy of Management Executive* 13 (4), pp. 36–47.
- Grinyer, P.H. and Yasai-Ardekani, M. (1980) 'Dimensions of organizational structure: a critical replication,' *The Academy of Management Journal* 23 (2), pp. 405–421.
- Reimann, B.C. (1973) 'On the dimensions of bureaucratic structure: an empirical reappraisal,' *Administrative Science Quarterly* 18 (4), pp. 462–476.
- Remland, M.S. (2004) *Nonverbal Communication in Everyday Life*. Boston: Houghton Mifflin.
- Robbins, P.S. (1994). *Management, 4<sup>th</sup> Edition*. Englewood Cliffs, NJ: Prentice-Hall.
- Sited in [<http://journals.iss.org/index.php/proceedings50th/article/view/308/92> accessed at 13:35 on 26/06/2009]
- Robinson, S.A. (1990) 'Improving organisational communication: developing a management strategy,' *International Journal of Public Sector Management* 3 (1), pp. 25–35.
- Samovar, L. and Porter, R. (2004) *Communication between Cultures*. Belmont, CA: Wadsworth Publishing Co.
- Twati, J.M. (2006) *Societal and Organisational Culture and the Adoption of Management Information Systems in Arab Countries*. [Phd thesis] Faculty of Griffth Business school, Brisbane, Australia.
- Udall, R. and Udall, S. (1979) *People and Communication*, Amersham, UK: Hulton Educational Press.
- Van den Ven, A.H. (1976) 'A framework for organization assessment,' *The Academy of Management Review* 1 (1), pp. 64–78.
- Wray, A. and Grace, G.W. (2007) 'The consequences of talking to strangers: evolutionary corollaries of socio-cultural influences on linguistic form,' *Lingua* 1 (17), pp. 543–578.

# E-Learning Adoption in Organisations: A Review of Existing Models and Practices

**Ugochukwu C. Digwo, Dumebi O. Oderinde and Mildred Brown-Houston**, *Business Logistics, Innovation and Systems Research Centre, University of Bolton*

## **Abstract**

This paper is an analytical review of the journey so far on e-Learning adoption by organisations, with a focus on identifying a model-based approach from existing research, and identifiable success factors. Every organisation needs to assess and identify where employee performance sits in their balanced scorecard in order to allocate sufficient resources, planning and strategy development for optimal success. This paper reviews the major trends in e-Learning for organisation while identifying critical elements that should be incorporated into every strategy framework or model developed. It discusses the need to adopt a pre-implementation analysis based on identifying organisational requirements, objectives and learning goals. There are current practices among leading organisations that are worth adopting as they are proven to be effective and have addressed some of the challenges faced by organisations adopting e-Learning. This paper is not a field research using any methodology, but simply a review of e-Learning models that best describes learning needs of both learners and organisations.

**Keywords:** e-Learning, e-Learning Models, Best Practices, Organizational e-Learning, Blended Learning.

## **1.0 Introduction**

As business cycle times compress and business processes become even more dynamic, workers from the plant floor to the executive offices require new skills almost weekly (Minton, 2000), and with demands on an employee's time seemingly overwhelming, new tools for delivering knowledge are needed. The need for employees to update their skills at the rate of technological advancement creates a sense of urgency for organisations to adapt to other learning channels. The mantra "fast-better-cheaper" was applied to learning strategies and one method that rapidly gained popularity was e-Learning (Monro, 2005). For this study, e-Learning is defined as the delivery of content, training and knowledge through electronic mediums. These mediums are in various formats, including, in current trends, audio, video, simulation, text-based, virtual classrooms and the mobile digital formats. This form of learning is innovative, accessible and flexible, among other characteristics. E-Learning as the transfer of knowledge through technology has become very important (Kathawala & Wilgen, 2004) with an increasing number of organisations adopting e-Learning as a strategy with each passing year. One of the goals and relative challenges of any human resources department is how to improve skills of employees so they are most efficient in performing their jobs. This is an evidential requirement for organisations to sustain strategic advantages in the business community in which they function. The lure of this learning strategy has been due to several identified benefits. Some of these benefits are listed by Kathawala & Wilgen (2004), building on the work of numerous authors who will be mentioned below. There are identifiable benefits in implementing e-Learning, which does not just include improving the performances of employees and learners, but benefits overall organisational performance. Hammer and Champy, in their book *Reengineering the Corporation*, groups success criteria to justify e-Learning investments into four measures; namely cost, quality, service and speed. These four groups are broken down into several units used to identify business benefits of e-Learning.

## **2.0 Identifiable benefits of e-Learning**

### *2.1 Cost-effectiveness*

A lower budgetary investment in e-Learning has been identified as one of the most quoted important advantages. E-Learning training delivery delivers savings on the huge expenses involved in training employees for many organizations. The main reason is that it eliminates or reduces travelling expenses, travelling time, instructors' pay, course delivery costs and maintenance over a period of two to three years (Kathawala & Wilgen, 2004). IBM estimated that it saved \$200 million in 1999, while providing five times the learning at one-third of the cost of previous methods (Strother, 2002 in Kathawala & Wilgen, 2004).

### *2.2 Business responsiveness*

Business practices and processes change rapidly within the business community, and organizations must be dynamic and flexible to manage these changes as a means of staying afloat. Business responsiveness is measurable by the rate of adaptability to change in the organization via the workforce. Having an e-Learning portal becomes critical for businesses when all employees need immediate awareness of new processes and adapt simultaneously for business to continue as usual.

### *2.3 Productivity improvements*

E-Learning, like training in general, can result in improved employee effectiveness. This is especially true when e-Learning is used to support on-the-job training. The opportunity to learn binds employees to the organization, which helps to decrease the employee turnover as a source of costs (Leary, 2001; Phillips, 2002; Oakes, 2003 in Kathawala & Wilgen, 2004). Keeping the employee turnover down helps organisations reduce knowledge loss. An e-Learning benchmark survey carried out by SkillSoft in 2007 within 16 European/Global status organizations, showed proof of increased productivity when e-Learning was deployed on the job. Most of the areas highlighted with increased levels of productivity include sales, customer service, and IT. There were areas of management and leadership skills that were also significantly improved, and this created an awareness of the value and prospects e-Learning holds for growing organizations. Other benefits identified by Kathawala & Wilgen (2004) include IT skills that resulted in an increase in overall employee performance level.

### *2.4 Customized and personalized learning*

This benefit is essential as it allows users to define their learning styles, when and where they want to learn, and to set the learning pace to one that meets individual needs. The learning content for every learner remains the same, but the method of communication via different learning formats to suit learners' preferences, makes learning easy. Technologies in use today include multimedia tools such as video conferencing and groupware. These tools incorporate video and audio facilities, as well as asynchronous tools that can provide anonymity in group-based learning. Learning can also be enhanced when it is made available uninterrupted via other tools including personal and mobile laptops, podcasts, webcasts, or MP3 players. A fraction of employees, 1.8% from the industries surveyed by SkillSoft suggests the introduction of learning through computer games. Though this is a contentious idea, it may be practicable when considering engineering and IT personnel who are more inclined to the technicalities of computing.

### *2.5 Faster learning and better retention*

Individuals learn more when learning is informal and in an informal setting. Learners consume larger volumes of information through on-the-job learning. Unlike traditional learning methods, e-Learning provides the user with the opportunity to recap on training wherever and whenever. The Research Institute of America found that, in traditional courses, by the second day after a class the retention rate of the studied material is 33 per cent, and at the end of three weeks, only 15 per cent of the knowledge is retained (Leary, 2001 in Kathawala & Wilgen, 2004). E-Learning in contrast leads to more control over the learning process, and a 60 per cent faster learning curve compared to instructor-led courses.

### *2.6 Content consistency and dependability*

Organizations that have global aims and international offices need uniformity and standardization for deploying important information across its value chain. E-Learning could serve as vehicle for delivering this consistency. E-Learning content can be enhanced with an effective knowledge management system and updated immediately and regularly, making it more accurate and reusable. All participants in a training programme can have access to the same material, which is dependable, thereby making it easier to ensure that the organization maintains thorough standards and consistent quality.

### *2.7 Universality*

There have been concerns in recent years about integrating differences in internet protocols, operating platforms and browsers, but this is ebbing away gradually with technologies that are more complex and with interoperable interfaces built into learning management systems. With technologies advancing, the concerns faced today regarding e-Learning programs are about the strict relevance of materials and content in these programs.

### *2.8 Availability*

E-Learning technologies provide a readily available resource for information and knowledge for learners. Employees are in constant need of information on-the-job and off-site, e-Learning provides a platform for learning and information transfer to take place even outside the confines of the office. Providing access to on-line books, reference ware and learner support, forms integral and essentials tools afforded by e-Learning. All e-Learning facility support is year round, twenty-four hours, seven days a week learning and access to information.

### *2.9 Interaction and knowledge sharing*

The Web today is a platform for communities of practice to share blogs and ideas on topics varying from professional practices to personal experiences. These tools are rich and can be used interactively for different purposes, which could be either for social networking or for knowledge insights and sharing. Several organisations go further towards integrating social networking tools such as messaging into their e-Learning programs, to help employees interact more and so share knowledge and ideas with other colleagues (Capuano et al., 2008). This aspect produces a wealth of information now converted into wikis and other online resources within the organization.

### *2.10 Scalability*

There is the potential for online courses to be scalable with respect to participants registering on a course. E-Learning programs can accommodate from 10 learners to 100 learners without many constraints, but only as long as there is standard infrastructure on the ground. The potential cost saving of running e-Learning programs, compared to other classroom-based training that takes the employee out of sight and away from the office, is potentially huge. The participants for e-Learning would only need to register for the course and plug their systems in or log on to join a class.

### *2.11 Source for knowledge management*

Knowledge Management, as defined by knowledge experts, captures information, data and tacit knowledge from within an organization, to improve business practices and overall performances. Technology on the other hand, enhances processes to provide employees with relevant and useful information, and best practices proven within the organization. The knowledge management philosophy is a vital aspect of e-Learning, as it affords knowledge experts the ability to share information from documents made available online with real-time communication and experiences, and to leverage organizational know-how with other employees through various media without physical inhibitors. AT&T Global Services Organization faced the challenge of how to retrieve market intelligence from their sales representatives in the field, bring them in for in-house training at the launch of a new product, as well as keep business information dynamic and relevant in the industry. They built an

Information and Knowledge Exchange (IKE) portal that could furnish their various teams who were constantly out in the field. The portal became the source of real-time information for new products among other things - market intelligence, new training available in real-time and best practices in the industry.

### *2.12 Corporate Leverage of Investments*

More managers and heads of departments are beginning to seek new opportunities for leveraging investments on local intranets for e-Learning. At the initial stage of introducing technology into business processes and internal activities within the organization, there were many failures of such IT projects characterized by huge spending. Organizational e-Learning has become an emerging tool within an organizational intranet and huge returns on investments have characterised the training embarked on by employees. Stakeholders consult e-Learning professionals to help streamline internal operations and transform the learning culture in the organization with less cost, richer learning content and a wider reach.

### *2.13 Private and Risk-free*

"Students taking an online course enter a risk-free environment in which they can try new things and make mistakes without exposing themselves" (Urdu and Weggen, 2000 in Kathawala & Wilgen, 2004). It affords users with the opportunity to experiment without having to worry about other people's attention.

As with any new tool, there have been drawbacks to implementing e-Learning in organisations. Opinions regarding e-Learning adoption have not been unanimous; Munro, (2005) states that there are many people who still do not like to use electronic means of learning.

There are five top drawbacks and barriers in successfully implementing e-Learning within organizations as highlighted in the Towards Maturity Benchmarking report (Overton and Hill, 2009). A study which showed that in 2004 IT infrastructure was the number one barrier to implementing e-Learning; in 2006, the cost of set up channelled the course for e-Learning and posed as a top challenge for organizations; while in 2008, a new set of barriers emerged as stated below:

- i. Reluctance by staff to adopt new technology
- ii. Lack of knowledge about its potential use and implementation
- iii. Lack of skills amongst staff to implement and manage e-Learning

## **3.0 Drawbacks in Implementing e-Learning in Organizations**

This next section outlines some of the major drawbacks in implementing e-Learning in both small and large organisations. They are faced with these challenges at different levels in implementing and sustaining e-Learning, but most of the drawbacks listed in this paper are common to them all.

### *3.1 Incoherent Training*

One of the main issues some employees have found about e-Learning, is that training materials have little relevance to what jobs or roles they play. E-Learning course content needs to be parallel to business objectives and goals, in order to for employees to perceive the training as top priority. Designing e-Learning course content should emanate from identifying what the organizational business drivers are and what steps need to be taken to improve overall employee performance.

### *3.2 Management Vision*

The question asked is what is the organizational management's vision for training employees? It is a widely shared belief among HR personnel that training is as critical as information sharing for the corporate capability of an organization to flourish. Management need to identify the fore-frontiers of the business and keep them well informed, trained to close any skill gaps.



### *3.3 Takes Too Much Time*

Training time restrictions constitute one of the setbacks faced when employees embark on in-house training. Managers are not readily disposed to employees embarking on an online course while on their desk or during a job. Courses that are lengthy and not exclusively relevant to employees' current roles and tasks are viewed as time wasters. E-Learning demands that employees allocate time off their jobs to complete a course, and this has not fared well enough since employees could not find time to complete the training needed for a job. This resulted in complaints and setbacks in projects that could not be completed within schedule, and employees reacting negatively to inadequate managerial support. Dell Computer Corporation, a high-tech company, found that to create an e-Learning culture required allocating time for its sales representatives to be off the phones and taking courses online.

### *3.4 Lack of Employee Motivation*

Employees need to know how and why the training offered through e-Learning is relevant and consistent for strategic business goals. In this way, they will not feel coerced into e-Learning but see the need and appeal of it. Oftentimes, these training materials should also reflect current job best practices for standardization across industries such as accreditations and future prospects for promotion. T-Mobile attracted a high percentage of employees in training programmes by offering free communication packages when they registered.

In a study carried out by Overton and Hills (2009) for the 'Towards Maturity Learning Technologies Benchmark Report, the 'people factor', or employee, poses one of the major drawbacks in implementing e-Learning in organizations. Over 50% of the 261 respondents to the study cited reluctance by staff to adopt new technology, while 40% of the SkillSoft study, 2007 identified a lack of knowledge of the e-Learning program and a low rate of involvement or commitment.

### *3.5 Costs Too Much*

This is a recurring drawback amongst most organisations looking to adopt e-Learning capabilities. The initial cost of setting up an e-Learning environment involves deciding what learning software design to use. The costs of IT hardware, consultancy fees and staff re-training amongst others, tend to be discourage organisations in adopting e-Learning strategies.

### *3.6 Human-Computer Interface as a Factor*

In attempting the e-Learning approach, some learners seem subtly menaced by the shift from the conventional face-to-face learning mode to the use of the Internet in learning. It is not an obvious option for learners who have not yet experienced or have not been prepared to use the new learning and knowledge transfer processes offered by a Web-based learning environment. On the other hand, replacing interpersonal relationships, typically occurring in conventional classroom-based settings, with interactions between the learner, tutor, and teacher in a Web-based learning environment, requires new skill and competence on the part of trainers, tutors, and authors of learning materials (Helic et al., 2005). The unsuitability of the e-Learning approach for another group not wholly disposed towards the web stems from different learning methods, approaches, and learning styles. The Visual-Auditory-Kinesthetic learning styles, also known as VAK model, provides a perspective of understanding and appreciating the different types of dominant thought life of an individual, and ways to improve their learning experience.

Several researchers argue that the way to improve learning is not by full adaptation of e-Learning, due to the tension it creates with respects to employee involvement, retention and motivation. These issues are being addressed by adopting "blended learning" for certain course training. Blended learning adopts a methodology of both e-Learning and a mix of other delivery methods. In summary, it combines traditional classroom learning methods and all its varieties with e-Learning in defined portions, designed to benefit from the advantages of both methods (Munro, 2005 & Kathawala & Wilgen, 2004). It involves delivering pure content

through e-Learning methods, while interpersonal skills development takes place within the classroom environment. The flexibility afforded by this method leverages both technology and economies of scale to capture a larger audience of employees retained and motivated for training.

### *3.7 Blended learning*

A major challenge faced by trainers is the issue of learning styles. It is widely recognized that different people learn differently: orally, visually, or using tactile or kinaesthetic forms (Piskurich, 2003). Most instructors try to varying the learning styles used in their classes, but in traditional learning mode the scales are tipped heavily towards those who have oral/auditory learning styles, and vice versa with pure e-Learning - several other learning styles could be disadvantaged.

Blended learning achieves a better result in addressing the issues around different learning styles. It provides instructors with a new package of tools that affords them the opportunity to blend more elements into both classroom and online PC based exercises. By offering learners a more diverse palette of learning options, they are more likely to adapt the learning programmes around their individual learning styles (Roshan, 2002; Piskurich, 2003). This decision on which mode to adopt e-Learning, fully or as part of a blended strategy, is one that organisational management needs to assess whilst considering an e-Learning adoption strategy. The next section addresses a major requirement of organisations considering e-Learning. Researchers propose that organisations conduct a detailed pre-implementation analysis that clearly defines the objectives, expectations, resources available (human, infrastructure and finance) to determine the best suited learning strategy and the mode of implementation as well as the organisation's strengths and weaknesses.

## **4.0 Pre-implementation Analysis**

It is also of the utmost importance to understand the individual business objectives and goals in a bid not to buy into existing and fashionable technologies, but to align business needs with e-Learning. Hartley (2000) advises individual firms to have a detailed framework of business goals and needs while considering e-Learning implementation. Researchers argue that before the decision to implement e-Learning as a learning strategy is taken, organisations need to conduct an internal assessment of their strengths, goals, needs and expectations (Roshan, 2002; Minton, 2000; Engelbrecht, 2003; Piskurich, 2003).

### *4.1 Proposed Pre-implementation Questions for e-Learning*

To effectively set up and administer an e-Learning program, there must be a parallel alignment of skill gaps and development needs with organizational goals and objectives. This assessment would ensure that the training is relevant, up-to-date, and scalable and that it addresses the specific business needs and objectives identified in the pre-implementation stage.

Some questions identified by research championed by Hartley, 2000 include:

*Functionality:* this deals with the practicality of the technology; what would the application do when it is installed?

*IT Specification:* this entails identifying your standard IT systems in the organization and their specifications. This is a bid to avoid complex interoperability and security issues when installed.

*User-interface specifications:* this helps you to identify how the interfaces work within the organization, as well as to determine what the end-screen would look like.

*Legacy system interface:* there is the issue of how the new application would interface successfully with other training software already in place. There is the question of support and standardization throughout the organization.

*User characteristics:* this requires initial identification of who would be using the application and how they would use it.

*Learning Content:* content must be measurable and performance-based to determine if it has any strategic value for the organization or not.

Some of the questions organisations need to answer at this stage include:

- What is it you want e-Learning to do for you?
- What business problems can it resolve?
- What should the end-result look like?
- Do you understand the changes e-Learning will bring to your organization?
- Is e-Learning part of your organization's integrated training strategy?
- Is there appropriate leadership throughout the organization to support e-Learning?
- Are the organizational support systems in place to sustain the adoption of e-Learning?
- Is your technology capable of delivering e-Learning predictably and effectively?
- Are learners prepared for distance learning?
- Do you have an overall Change Management plan in place to transition your organization to e-Learning?

#### *4.2 Proposed Pre-implementation Questions for Blended Learning*

Piskurich (2003) proposes that organisations that would rather adopt the blended learning strategy, or might be considering it, would have to consider an effective mix of learning styles.

To determine the most effective blended learning mix for their training projects, organisations will need to gather information regarding the following:

*Target Audience:* It is important to know as much as possible about the target audience when designing blended learning. Particularly, the familiarity and comfort level of the target audience regarding learning via computers will have to be determined.

*Skills/Content:* The type of skills and content learners need to master will greatly influence the delivery methods selected. Some skills, such as supervisory and communications skills require interaction with an instructor, while others like technical skills and basic cognitive material are best suited for a self-paced environment. By matching the content carefully with the methods available, instructors will ensure the effectiveness of the programme.

*Technical resources:* Blended learning relies, largely, on the technical resources of the organisation. Therefore, careful cataloguing of what the organization does and does not have will narrow down the choices available and avoid costly mistakes.

*Personnel resources:* The personnel resources available for development, delivery, and technical support will influence organisational learning choices. Training staff not experienced in e-Learning methods will have to be done through options that are easiest to design, roll out, such as on-line discussion, and chat groups. Another alternative would be to use third-party vendors for some aspects of the project.

*Budget/time constraints:* In the real world, there are always budget and time constraints that organisations have to consider in choosing their blended learning mix. Organisations have to remember to factor in maintenance costs and timeframes as part of the constraints they will need to manage.

This then leads to the next consideration identified in the pre-implementation analysis, the e-Learning model approach best suited to the organisation because there are several proposed models that has been applied successfully in different industries. These recommended models are reviewed in the next section.

## 5.0 A Model-based Approach for e-Learning

Adopting a model-based approach for e-Learning involves finding the right models, concentrating on organisational core competencies, targeting the right group and designing the right product for deployment. There are different models of e-Learning more suitable for different businesses to adopt either for medium or long-term success sustainability. Current assessment of e-Learning still indicates that proportional needs of learners are not fully realised in a formal learning environment compared to any informal method. There is the urgent and compelling need for management to approach e-Learning to meet specific needs in the organisation and to improve the performances of new graduates. The National Audit Office for the UK reported in 2005 that despite large budgetary expense on education, six per cent of employees have skill shortages and twenty per cent have a skills gap, equivalent to ten billion pounds worth of losses in revenue (National Audit Office, 2006). The need for business-process models for e-Learning is becoming a more dominant issue for organisations. As noted by LeyKing et al. (2007), employees are the intellectual assets of any organization and thus, a continuous development of skills and competencies is inevitable for optimization of processes. A correlation between business-process models, business process, and the learning process has to be defined. This is to provide advantages for organizations and learners by a "Process-Oriented Learning and Information Exchange", PROLIX, a project which began on 1st December 2005 under the European Framework Programme for Research and Technological Development. Its main objective is to identify complex business processes and identify individual and organizational learning goals to be aligned for efficiency and effectiveness. This model supports an all-round approach for individual learning lifecycles (Martin & Wolpers, 2005).

Competence-driven content development emphasizes the suitability of e-Learning content designed to close any competency gaps and requirements identified by the previous process. Course content and learning goals become relevant, suit a learner's immediate need to carry out his job effectively, and the learning process becomes coordinated with organizational goals. This platform authors a new chapter for organizational own learning management systems that render on-time learning tailored to individual profiles (Zimmermann et al. 2005). 'EXPLAIN' is a content development project funded by the Federal Ministry of Economy and Technology in Germany. The focus was to develop a technology-enhanced environment for organizations to build learning objects.

Implementing business-integrated learning management serves as a strategic tool for smooth organizational performance where individual profiles of employees are enhanced based on process competencies, with a view to close skill gaps and be ready for anticipated change in the organizational knowledge mapping structure.

### 5.2 E-Learning models

E-Learning models are attempts to develop frameworks to address the concerns of the learner and challenges presented by the technology, so that online learning can take place effectively. In the strategic planning process, these models provide useful tools for evaluating existing e-Learning initiatives or determining critical success factors (Engelbrecht, 2003). By 2008, the e-Learning market should have more than doubled rising to \$13.5 billion in the United States and \$21 billion globally, according to International Data Group Inc. subsidiary International Data Corp. (IDC), a research firm that uses almost 800 analysts in 50 countries to track technology and industry trends (Tucker, 2005). A large number of organisations across diverse industries worldwide are integrating e-Learning into their learning strategy and several models have been designed. Two of these models have been selected in this section and are based on the different components included. These tools are used to enhance learning, keeping the learner and organisational needs in focus.

### 5.3 The Wyeth e-Learning Model

This e-Learning model gives a description of what it entails in aligning organisational learning with business objectives and strategies for success. Every good framework should provide multiple channels for learning and support in order to create flexibility and unlimited access to information. This model consists of two major components:

(i) *Learning and Performance Support*: the first component provides employees with various learning opportunities based on business needs and learning styles. It utilises the blended learning style by providing the organisation with the opportunity to combine traditional learning methods with that of e-Learning. These opportunities include learning through instructional methods such as Web-based training (self-study), or instructor-led training that can take place either in traditional classes or in virtual classrooms through synchronous Web-based access. Learners are able to gain tacit knowledge through interaction with business experts and peers who are experiencing or have experienced similar learning processes in real project based work environments; or similarly, learning through collaboration with other company employees worldwide, or with business partners, external research institutes, and academic experts. Finally, there is learning through knowledge assets of the company, such as internal or external research findings, articles and other documents, best practices templates, or any other company lessons from the past.

(ii) *Competency Management*: Roshan (2002) states that this component provides a learner-centric and user defined learning environment. This creates opportunities for learners to customise their performance improvement, with an emphasis on each individual employee's career path and development, performance needs, desired outcomes, and personal interest. Of great advantage to management are the performance administration capabilities, such as tracking and monitoring of learning processes by training administrators or supervisors, who might suggest attendance at classroom instruction, seminars, and conferences, or call for reporting and documenting of the employee's learning achievements.

Figure 2 shows the two components of the e-Learning model are tightly linked and interdependent; to provide a continuous and active learning environment that supports the company's business objective (Roshan, 002).

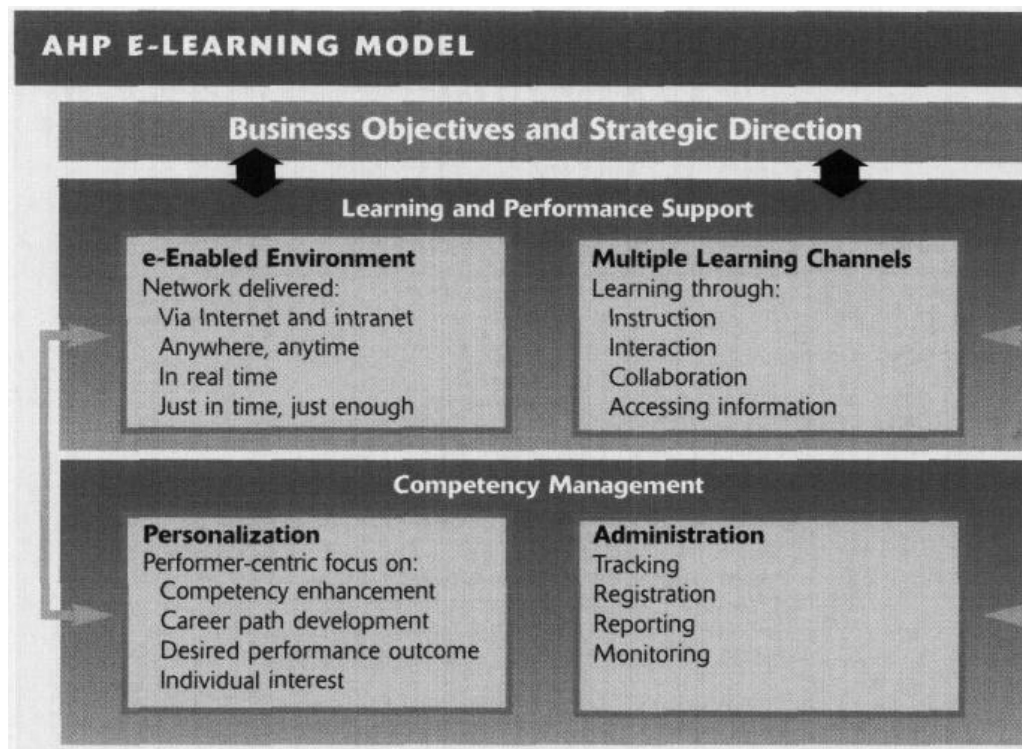


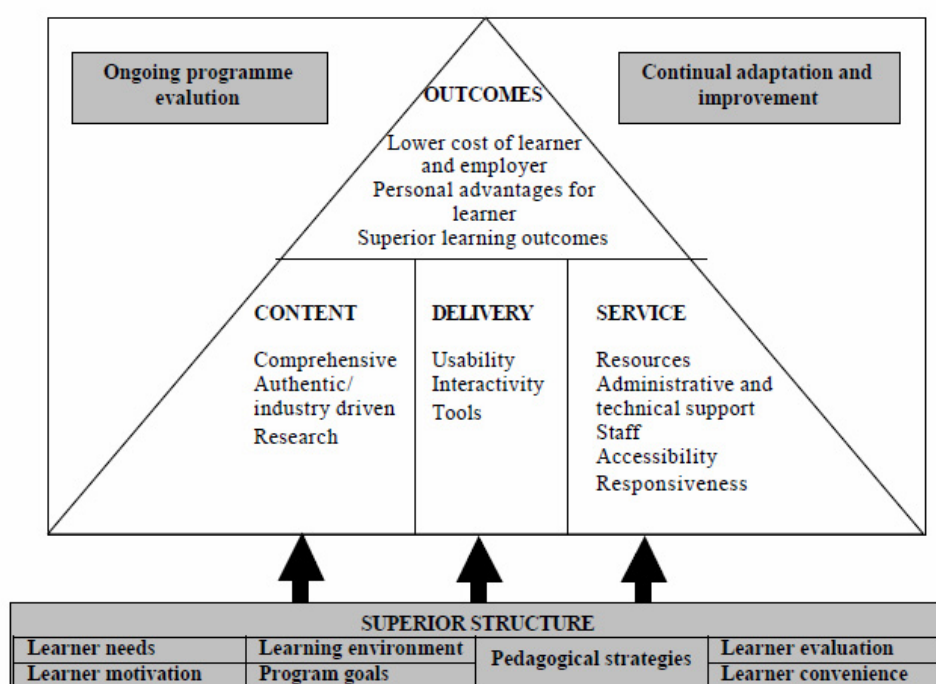
Figure 2 The Wyeth E-Learning Model (Roshan, 2002).

The next model could be described as a further step from the Wyeth e-Learning Model to a more detailed structure for designing an e-Learning strategy. The model also highlights the

importance of realising the changing needs of learners and their employers, and the pedagogical changes that could be made to content and services in order to meet these needs (Engelbrecht, 2003).

### 5.3 The Demand-Driven Learning Model

This model was developed in Canada as a collaborative effort between academics and experts from private and public industries (MacDonald et al., 2001). The model sees technology as support or a tool to achieve the desired learning outcomes in a cost-effective way. The primary purpose of the model is to encourage academics (and organisations) to take a proactive role in the development and use of technology in the teaching/training process. It emphasises the three major consumer demands: high quality content, delivery, and service (Engelbrecht, 2003). Content should be comprehensive, authentic and researched. Delivery is web-based, and the interface of e-Learning programmes should be user-friendly with communication tools to support interactivity. Service should include resources necessary for learning and administrative/technical support.



**Figure 3.** The Demand-Driven Learning Model, (MacDonald et al 2001).

It is important to have a long-term planning vision for e-Learning. Priority course content should fulfil high-profile critical needs within the organisation. Content should be delivered in small chunks for effectiveness. Courses should portray segments of jobs handled by a particular division; this way, the learning will become relevant to a team's immediate need and create motivation vital for business performance and other courses once successful. The Grant Thornton University follows this concept of web-based tailored content for learners. Bob Dean, the chief learning officer, held onto the philosophy that daily learning at work is the key to competitiveness and profit for every organization. He created a theory that the course contents for employees attending or accessing the university must follow a customized learning path aimed at addressing business needs and providing solutions for each business unit. Therefore, no two departments were undergoing the same online learning. Managers would work with employees by accessing feedback on individual performances and identifying needed competencies and skills for the department and relevant training courses. They would also choose the most appropriate method of delivery for each particular course topic that was not limited to self-paced study, but included live virtual classrooms, workgroups, question-and-answer sessions, and case studies.

The model at the Grant Thornton University is a blended learning model that suited most of the company's learner types. It eliminated the antagonism developed by employees towards e-Learning with respect to its relevance and suitability. It created the perfect environment for

learners to liaise with their peers and work experts who were usually in separate groups. This way, they did not feel abandoned in learning, or uncomfortable learning online.

Current trends of e-Learning seem to evolve continuously around the objectives of promoting a flexible and dynamic method of learning. A survey carried out by the Guild Research in 2004 shows that 48% of learning takes place in an informal way or among peers; while formal training with learning objectives takes a portion of 23% and other learning 29% via learning on the job. Conrad and Donaldson (2004) hold strongly to the thought that more is being achieved by e-Learning through the increased level of creativity and critical thinking. Two key issues addressed continually in e-Learning development for the way forward include the need for new methods and skills for collaboration, and the need for effective online assessment. These will become fundamental packages necessary in designing e-Learning technologies for the future:

- a) Collaboration: some learners tend to be more attracted to learning when it is conducted through visual representations. This is one of the issues addressed in learning styles as a guide to building e-Learning modules and course content. Jitendra, (2002) calls for more active online learners, resulting in higher satisfaction levels when visuals representations are used in communicating course content.
- b) E-assessment: The use of quality and benchmarking tools for e-assessment has grown considerably and stands to validate e-Learning as a viable learning method. Mason (2002) suggests stronger measures be adopted in e-assessment to create valid yet flexible and authentic assessments that cater to various learning styles.

#### *5.4 Examples of Demand-driven e-Learning Models and Best Practices*

Today, organizations are faced incessantly with the challenge of providing training and workshops for employees in line with roles and responsibilities for high effectiveness on the job. More and more organizations use Learning Management Systems to integrate learning for employees at work and attempt to capture and assess overall performance. Sixty percent of North American IT and business professionals confirmed, in a survey, that it is extremely important for content delivered through these learning systems to be in the context of the business process in which they are involved.

There are downsides and repercussions when management in organisations do not consider e-Learning components thoroughly, as well as when it is inaccurately prioritised. High personnel turnover in the automobile industry has reached a peak of 75%, and this has created a hole in delivering best customer experiences to car buyers. On-site sales representatives need to have quick access to newly launched products, sale-points and services, as well as a working knowledge - especially for cars with 'high tech' specifications. Forrester (2005) suggests an integrated e-Learning portal with training content vendors like Global Knowledge Network and Thomson NETg and SkillSoft.

There have been good examples of organisations that have executed e-Learning strategies which aligned with their organisational goals and objectives, and that have been successful. Cable & Wireless, a British telecommunications company, began reinventing its technological business strategy to be more service-driven by focussing on its largest group of users of telecommunications globally, and reducing complexity in products, systems, and processes. SkillSoft developed a one-stop shop learning system, iLeARN, in which employees can pick specific e-Learning options tailored to related tasks at their time of need. It provides an increased level of access and flexibility for learning, which is fundamental in meeting business needs both quickly and economically. The learning courses and content are created on short timescales due to the rapid changes in business processes and technology. This supports the overall business transformation, where e-Learning and its technologies become integral and critical components of the business. Today, Cable & Wireless maintains a successful business strategy in e-Learning (Becta, 2008).

Nike Corporation, a major sportswear and equipment supplier with headquarters in the United States, used specialized online training to drive a 2% increase in sales and so did Best Buy, an American electronic retail store. Its slow moving products increased in sales by 20% after online training was adopted (Bünger et al., 2005). Nike's e-Learning portal integration allows employees to choose courses and modules aligned with organizational goals, but with reference to regional activities. These include online discourse that surrounds the unit, task-specific and soft skill development, which could be virtual classrooms, training sessions with tutors on particular topics they need, unscheduled practical sessions, workshops or audio lectures. Materials from experts' repositories, such as the technicians' message boards and other internal forms of communication, can be added for other colleagues and teams to use. The method used by employees to glean information and gain knowledge becomes broad and easily accessible. Their e-Learning experience becomes a personal online training portfolio that goes with them as they move on through other roles and regions in the organisation. The portal is accessible by management to monitor and assess employee progress, content case study, and reusability. This was also typical of GM's DealerWorld Workbench and IBM's Sales Compass Portal. IBM's Sales Compass Portal provides on-the-job training on its Signature Selling Method (SSM) to 60,000 sales representatives and consultants worldwide.

### *5.5 The IBM Sales Compass Portal*

The goal for the Signature Selling Method is to keep IBM's market valued skills responsive to changing markets, industry trends and client demands (IBM, 2004). Four major objectives were set out towards meeting these changes:

1. To re-skill the workforce quickly on limited schedules
2. To do so on cost-effective and limited budgets
3. To continuously adapt to changing requirements
4. To globally address and meet the training needs of a dispersed and mobile workforce.

In summary, with evident proofs of these best practices and models being turned into business results, more organisations are beginning to reassess and evaluate their potential e-Learning programs. Many organisations are developing a balanced mix of both formal and informal content for employees, and view e-Learning and learning investments as major projects and top priorities. The agenda of positioning business objectives at strategic points helps create value for learning and for economical advantage. We can also say that every good model adopted has been pruned thoroughly to meet only essential needs of learners.

## **6.0 Conclusion**

We can conclusively say that e-Learning in organisations is very much a part of organisational strategy in present times and looks to be a trend that most, if not all, will be adopting in the near future. In this paper, we can say that existing research show that there are drawbacks and benefits to e- adopting e-Learning practices; this should serve as a set of guidelines to organisations seeking to join the bandwagon, or organisations presently reviewing existing possibilities. E-Learning should not necessarily be a stand-alone strategy or an alternative to traditional learning methods. Blended learning provides the opportunity for organisations to combine the best feature of both learning methods. Everything starts with a first step which, in this case, is the pre-implementation analysis aimed at assessing an organisation's ability to properly implement and support an e-Learning strategy; a look before the jump. Finally, e-



Learning models are attempts to develop frameworks to address the concerns of the learner and the challenges presented by the technology, so that online learning can take place effectively. A framework that incorporates all these desired elements guides a successful strategy.

## References

- Becta, A. (2008), 'Case Study: Cable & Wireless'. Available from <http://www.nextgenerationlearning.org.uk/en/Employers/Resources/Case-study-Cable-Wireless/>. [Accessed February 26, 2009.]
- Capuano, N; Gaeta, M; Ritrovato, P; Salerno, S. (2008), 'How to integrate technology-enhanced learning with business process management' *Journal of Knowledge Management* 12 (6), pp. 56-71
- Conrad, R. M., & Donaldson, J. A. (2004), *Engaging the Online Learner*, San Francisco: Josey-Bass
- Engelbrecht, E. (2003), 'A look at e-Learning models: investigating their value for developing an e-Learning strategy Bureau for Learning Development, Unisa' *Progressio* 25(2), pp.38-47. Available from [http://www.unisa.ac.za/contents/faculties/service\\_dept/bld/progressio/docs/engelbrecht.pdf](http://www.unisa.ac.za/contents/faculties/service_dept/bld/progressio/docs/engelbrecht.pdf)
- Hartley, D. (2000), 'All aboard the e-Learning train. Training and Development; ABI/INFORM' *Global*, 54 (7), p.37
- Helic, D.; Krottmaier, H.; Maurer, H.; Scerbakov, N. (2005), 'Enabling Project-Based Learning in WBT Systems', *International Journal on eLearning*, 4 (4) 4, pp. 445-461.
- IBM Learning Solutions 2004. 'IBM's learning transformation story'. *IBM Internal Research*. Available from <https://www.ibm.com>. [Accessed February 03, 2009]
- Hammer, M. and Champy, J. (1993), *Reengineering the corporation – manifesto for business revolution*, New York: Harper Business
- Jitendra, A. (2002), 'Teaching students' math problem-solving through graphic representations', *Teaching Exceptional Children*, 34 (4) pp. 34-38
- Kathawala, Y & Wilgen, A. (2004). 'E-Learning: evaluation from an organization's perspective'. *Training & Management Development Methods*, 18, (4), pp.501-513.
- LeyKing, K; Chikova, P. & Loos, P. (2007), 'Competency and Process-Driven e-Learning: A Model-Based Approach', *The Electronic Journal of e-Learning*, 5 (3), pp. 183-194.
- MacDonald, C J; Stodel, E J; Farres, L G; Breithaupt, K; Gabriel, M. A. (2001), 'The demand-driven learning model: a framework for web-based learning', *The Internet and Higher Education*, 4, pp.9-30.
- Martin, G. & Wolpers, M. (2005), 'Process-driven Learning and Knowledge Environments', in M. E. Auer, U. Auer and R. Mittermeir (eds.): *Ambient and Mobile Learning. Proceedings of the Interactive Computer Aided Learning Conference (ICL) 2005*, Carinthia Tech Institute (School of Electronics), Villach.
- Mason, R. (2002), 'Rethinking assessment for the online environment' In C. Vrasidas & G. V. Glass (eds.), *Distance education and distributed learning*, Greenwich, CT: Information Age, pp. 57-74
- Minton, M.C. (2000), 'Is Your Organization Ready for E-Learning? - Seven key questions you need to answer', *Communication Project Magazine*, Vol 3.1 (Summer 2000).
- Munro, R.A. (2005), 'Understanding the Buzz Around E-Learning: Searching for Faster/Better/Cheaper Learning - Effectiveness of E-Learning Techniques', *ASQ World Conference on Quality and Improvement Proceedings* 59, pp. 131-143. Available from <http://www.proquest.com/>
- National Audit Office, (2006). National Audit Office Report: 2006. Available from <http://www.nao.org.uk> [Accessed January 28, 2009]

- Overton, L, & Hills, H. (2009), *Driving Business Benefits. Towards Maturity Learning Technologies Benchmark Report*. Available from <http://www.towardsmaturity.org> [Accessed February 10, 2009]
- Piskurich, G. (ed.) (2003), *The AMA Handbook of E-Learning*, American Management Association. New York: AMACOM
- Roshan, V. (2002), 'How one corporation got a charge out of e-Learning', *Medical Marketing and Media*, February 1, 46-52. Available from <http://www.proquest.com/> [Accessed January 28, 2009].
- Tucker, M. A. (2005), 'E-Learning Evolves', *HRMagazine*, October 1, pp.74-78. Available from <http://www.proquest.com/> [Accessed February 28, 2009]
- Zimmermann, V; Bergenthal, K.; Chikova, P.; Hinz, D.; Lehmann, L.; LeyKing, K.; Martin, G.; & Rensing, C. (2005): 'Authoring Management Platform EXPLAIN', in *Ariadne PROLEARN Workshop*, TU Berlin, pp. 1-7

# E-business Adoption as a Form of Management Strategy: An e-Retailing Perspective

**Ugochukwu C. Digwo and Mildred Brown-Houston**, *Business Logistics, Innovation and Systems Research Centre, University of Bolton*

## Abstract

This research focuses on the increasing importance of e-retailing as a major mode of retailing, which retailers must integrate into their management strategy, if they are to maintain or gain competitive advantage in the industry. The study examines the benefits and challenges that accompany e-retail adoption or non-adoption by retail organisations. To this end, a survey was conducted to determine the views and preferences of customers and retailers in the researchers' community regarding the adoption or non-adoption of e-retailing. The results of the research are positioned as a frame of reference for designing organisational management strategies.

**Keywords:** e-retailing, management strategy, e-business

## 1.0 Introduction

E-business, as it is known today, arose through the widespread use of the internet as a platform for inter-organisational systems (IOS) in the late 1990s. Zwass (1997) observes that in recent years, the use of the internet has increased the sharing of business information and built business relationships. It has also enhanced business transactions by means of telecommunications networks and has been particularly significant for developments in the operation and strategic management of supply chains and global networks. In support of this widely held view, Evans and Wurster (2000) observe that the rise of the internet as a communication channel (and its supporting systems and software) has changed the economics of information. It has given rise to new opportunities and new forms of affiliation and transaction amongst organisations.

E-business has therefore radically transformed virtually all aspects of organisational life, including commerce (international, national, and local), procurement, manufacturing, customer service, and business alliances.

One of the most notable areas of retailing transformed by e-business is supply chain management, which includes processes such as customer relationship management, customer service, demand management, order management, production and material flows, and purchasing (Cagliano et al. (2003). E-business internet tools can be classified as follows: e-commerce, which provides support to sales, distribution, and customer service processes; e-procurement provides support to sourcing, procurement, tendering, and order-fulfilment processes; and e-manufacturing, which provides support for demand and capacity planning, forecasting, and internal supply chain integration. Lee and Whang (2001; cited in Cagliano et al., 2003:23) identifies a fourth category, e-collaboration., which refers to the use of the internet as a means of strengthening relationships along the supply chain through the exchange of data and collaborative decision making.

Frohlich and Westbrook (2002) point to the additional phenomena of low integration and demand integration (which is typical of e-commerce); supply integration (which is typical of e-procurement); and demand chain management integration (which is the joint application of the previous two strategies).

E-business deployment means different things to different parties. For instance, in the view of industrial marketers, e-business has triggered a growth in interest in network (rather than dyadic) levels of activity, which concentrates decision making on issues of supply chain optimisation (Simon R. Croom, 2005). In this context, management consultants Accenture conducted an e-fulfilment survey in 2001. E-fulfilment refers to the ability to completely satisfy customer needs across the supply chain—in either the business-to-business (B2B) or the

business-to-consumer (B2C) channel—with emerging web-based technologies in the most cost efficient manner possible. Accenture’s white paper showed that over 72% of the respondents—which included over 100 logistics firms from over seven European countries, including UK—saw e-business adoption as vital to their continued existence, (Accenture, 2001).

## 2.0 Mitigating factors governing e-business adoption

Notwithstanding its enormous popularity and growth, several factors govern the adoption of e-business as a management strategy. As with every strategy, e-business entails attendant risks, such as security, dehumanisation of the transaction process, and customer dissatisfaction. Bocij et al. (2003) includes some of the following as vital factors governing adoption:

- 1) *Cost of access*: The cost of accessing e-business can be a barrier for those who do not own a personal computer, which is indeed a major expenditure for many households. Computer ownership also entails attendant expenses, such as an ISP connection, telephone, and cable charges. In view of these costs, free access would certainly increase adoption and usage.
- 2) *Value proposition*: Before an individual or an organisation can be persuaded to invest or participate in any form of e-business, the question of what they are to gain must first be addressed. Customers need to perceive a need to be online: What can the internet offer that other media cannot? Would-be adopters must be offered incentives—for example, access to more supplier information and lower prices when conducting business online.
- 3) *Ease of use*: Chaffey et al. (2003) describe this factor as the ease of connection to the internet and the ease of use when connected. In this context, human-computer interaction (HCI) and web navigation are vital; consumers must be able to navigate a website and locate required information easily. In addition, the aesthetics of the site must be appealing.
- 4) *Security*: This is a widespread concern amongst online shoppers, many of whom view themselves as being at risk of having their personal details, such as credit card details, stolen by internet bandits (Chaffey et al. 2003).
- 5) *Fear of the unknown*: Many would-be adopters have a general fear of new technology. This is compounded by negative press about potential risks such as identity theft, hacking, fraud, and privacy infringements.

## 3.0 Challenges to e-business adoption

Many companies are starting to use the internet to communicate with both their customers and suppliers, thus creating new digital electronic commerce networks that bypass traditional distribution channels. The use of internet (or e-business) applications extends far beyond just the creation of digital networks; a lot of companies are using these applications to streamline their internal business processes as well. Laudon and Laudon (2003) observe that digitally enabling business processes and relationships with other organisations can help companies achieve new levels of competitiveness and efficiency, but they identify the following as possible challenges that management will face:

- 1) *Digitally enabling the enterprise requires a complete change of mindset*: As one of their most crucial challenges, digital firms will need to acquire new organisational designs and management processes. Companies must examine and perhaps redesign entire business processes rather than try to graft new technologies on existing business practices, if they are to effectively use e-business applications for organisational coordination, collaboration, and electronic commerce successfully. In the words of Laudon and Laudon (2003), “companies must consider a different organisational

structure, changes in organisational culture, a different support structure for information systems, different procedures for managing employees and networked processing functions and perhaps a different business strategy.”

- 2) *Finding a successful internet business model:* As the dot.com era arose, companies all over the world raced to put up websites in the hope of reaping mega profits through e-commerce. However, many of these e-commerce websites have yet to turn a profit or to make any difference in the business’s sales and marketing efforts. The promises of instant profits and new markets have failed to materialise. In a nutshell, internet technology alone is not a substitute for an effective business strategy; businesses need to do a careful evaluation of what they want, how e-business relates to their overall business strategy, and how they can create a genuinely workable business model using e-business applications.

#### **4.0 The e-(re)tailing perspective**

Short for “electronic retailing,” e-retailing, or e-tailing is the selling of retail goods on the internet. Coming into use as early as 1995, the term—an addendum to terms such as e-mail, e-business, and e-commerce—applies to business-to-consumer (B2C) transactions (Schappell, 2000; Jones et al., 2002). As early as 1997, the practice began to work for some major corporations and smaller entrepreneurs. In that year, Dell Computer reported multimillion-dollar orders taken at its website. The success of Amazon.com and eBay hastened the emergence of other e-retailing websites. Concerns about secure order-taking, previously a serious a topic of debate, rapidly receded. Also in 1997, the website Auto-by-Tel reported the sale of its millionth car, and CommerceNet/Nielsen Media reported that 10 million people had made purchases on the web and suggests that e-tailing would grow to \$37 billion in revenues by 2002.

E-tailing has resulted in the development of e-retailware—that is, software tools for creating online catalogues and managing e-tailing-related business. A new trend is the price comparison site, which can quickly compare prices from a number of different e-retailers and link you to them. Varying views exist on the future of e-retailing. While some analysts (e.g., Verdict, 2000; Foresight, 2000) have predicted substantial growth in e-retailing, others (such as Barsh et al., 2000; Edgecliffe-Johnson, 2000) have taken a much more measured, if not pessimistic, position.

The opportunities and challenges offered by the “electronic retail marketplace” are generating both excitement and concern within the business community. On the one hand, e-retailing offers companies the opportunity to communicate with existing and new markets in a sophisticated manner and to expand their business horizons to satisfy customer demand in global markets. On the other hand, there are concerns about the fundamental financial viability of e-retailing. In addition, developing brand awareness and value; designing attractive and effective websites; establishing reliable systems to handle and process orders; managing the mechanics of payment and delivery; and the security of financial transactions are potential problems for online traders. Peter Jones et al. (2002) delineate these challenges as follows:

- 1) Will traditional retailers with substantial capital investment in often expensive sites and properties see e-retailing as posing a major threat to their operations and profitability? If so, how can they respond quickly and effectively?
- 2) What of customers themselves? Will they take quickly, loyally, and in large numbers to electronic shopping, and how will they access e-retail sites?
- 3) How can e-retailers manage customers’ expectations that e-retail prices should be cheaper than those of conventional shops?
- 4) What consumer protections and guarantees will there be in these virtual markets, which seem to transcend conventional international boundaries and nation-state jurisdictions?

These crucial questions are being addressed by various researches on e-retailing. This field is still new and the likely rapid development of increasingly sophisticated interactive, multimedia communication technologies linked to the internet may continue to bring rapid and perhaps unforeseeable change to the retail marketplace and to shopping behaviour in general. This

research will seek to analyse the prospects of companies wishing to adopt e-business and make relevant recommendations.

Companies with an interest in e-retailing can be split into three groups:

- 1) New companies established to trade specifically and exclusively online; examples are Amazon and eBay, the world's biggest e-retailers.
- 2) Existing retailers that were established as traditional stores or mail order operations; examples include Tesco, Marks & Spencer, and Littlewoods.
- 3) Manufacturers who are attracted by the possibility of dealing directly with the ultimate consumers of their products. Dell is a notable example of this trend.

Amazon.com is perhaps the best known dedicated e-retailer in the UK. The company opened its virtual doors in 1995 with a mission to use the internet to transform book buying into the fastest, easiest and, most enjoyable shopping experience possible. Five years later the company had some 23 million customers in over 160 countries. Annual sales were running at \$2.3 billion, and its product range had been expanded to include over 18 million items in categories including CDs, DVDs, toys, videogames, electronic equipment, software, garden and patio, tools and hardware, kitchen, health and beauty, and home living. Through Amazon.com zShops any business or individual can sell virtually anything to Amazon's 23 million registered customers (Jones et al., 2002)

### **5.0 Expected benefits to be derived from e-retail adoption**

The opportunities and challenges offered by e-retailing have attracted considerable attention from the business community, government, and consumer groups. The business media and the general media have both identified a wide range of benefits to be gained by establishing e-retail formats. Jones et al. (2002) and Burgess (2006), enumerate these advantages as follows:

For e-retailers:

- 1) accessibility to a potentially global market 24 hours a day
- 2) reduced costs: advertising, support, and transaction expenses are less for electronic stores, making them cheaper to establish and maintain than traditional shops
- 3) the ability to make rapid changes in product specifications and prices
- 4) greater customer involvement in product and service innovation
- 5) perhaps most important of all, increased profitability.

For online shoppers:

- 1) ease and speed of purchase
- 2) the convenience of comparison shopping from, and delivery to, the customer's home
- 3) good quality and topical information on products and services
- 4) the availability of products or services that may not be obtainable from local conventional shops and businesses
- 5) potentially cheaper prices
- 6) online support for after-sales service

### **6.0 Challenges of e-retailing adoption**

Notwithstanding all of its potential benefits, most companies that have adopted e-retailing have encountered some obstacles. These challenges range from high initial implementation and maintenance costs to overestimating the benefits of the adoption.

Several researchers have developed theoretical frameworks for the adoption of e-retailing; however; putting these theories into practice has not always been easy. Amazon.com is currently listed as the forty-eighth most valuable brand in the world and acknowledged as a

leading global e-retailer. However, in the summer of 2000, the company admitted that sales growth had not been as strong as expected. Amazon's annual operating losses were running high; its long-term debt commitments and liabilities were \$2 billion; its share value was about 66 percent below its late 1999 peak; and many US analysts had ceased to recommend the company to its clients. In 2000, a number of high profile UK e-retailers faced collapse. Among these were Clickmango.com (health and beauty specialists) and Boo.com (specialising in fashion sportswear), which lost some \$100 million of venture capital in less than two years. These examples serve as pointed reminders of the fragility of the e-retail marketplace (Newbery, 2006; Jiang and Talaga, 2006).

A major challenge to creating trading profitability and removing debt burdens in the e-retailing industry is the ability of businesses to develop the marketplace and establish a large customer base quickly. To this end, Lastminute.com, for example, initiated a management strategy that stresses the importance of building brand loyalty and awareness, growing supplier relationships, developing its value proposition to customers, extending its marketing channels to embrace a wide range of electronic media and platforms, and early investment in leading technologies. At the same time, the company has been keen to impose stringent controls on its operating costs. In the first six months following its flotation on the stock market in March 2000, Lastminute.com did report operating losses but investment analysts were of the opinion that it would begin to deliver net profits by the end of 2002. More generally, e-retailers also need to address a number of business fundamentals if they are to grow in the e-retailing industry. In this context, some researchers (Trabold et al., 2006; Joia and Sanz, 2005; Merrilees and Miller, 2005; Maltz et al., 2004) have identified the following challenges:

- 1) Ensuring customer retention and transaction profitability
- 2) Establishing easily accessible and user friendly websites
- 3) Offering a range of high-quality products and services at competitive prices
- 4) Establishing speedy and reliable delivery systems
- 5) Providing information on quality assurance and after-sales service, as well as guarantees about the security of online financial transactions and the privacy of personal information
- 6) Creating a pleasurable shopping experience. Many people view the act of hands-on shopping as a socially valuable and personally relaxing experience (described by some as "retail therapy"); the virtual creation of such experiences may prove an elusive challenge.
- 7) Enabling logistics; search and inventory management; and physical delivery
- 8) Establishing customer loyalty and satisfaction

A recent research survey by the Taylor Nelson Sofres (TNS), one of the world's leading market information companies, identified some of the most common customers' complaints about their online e-retail patronage (see Table 1).

**Table 1.**

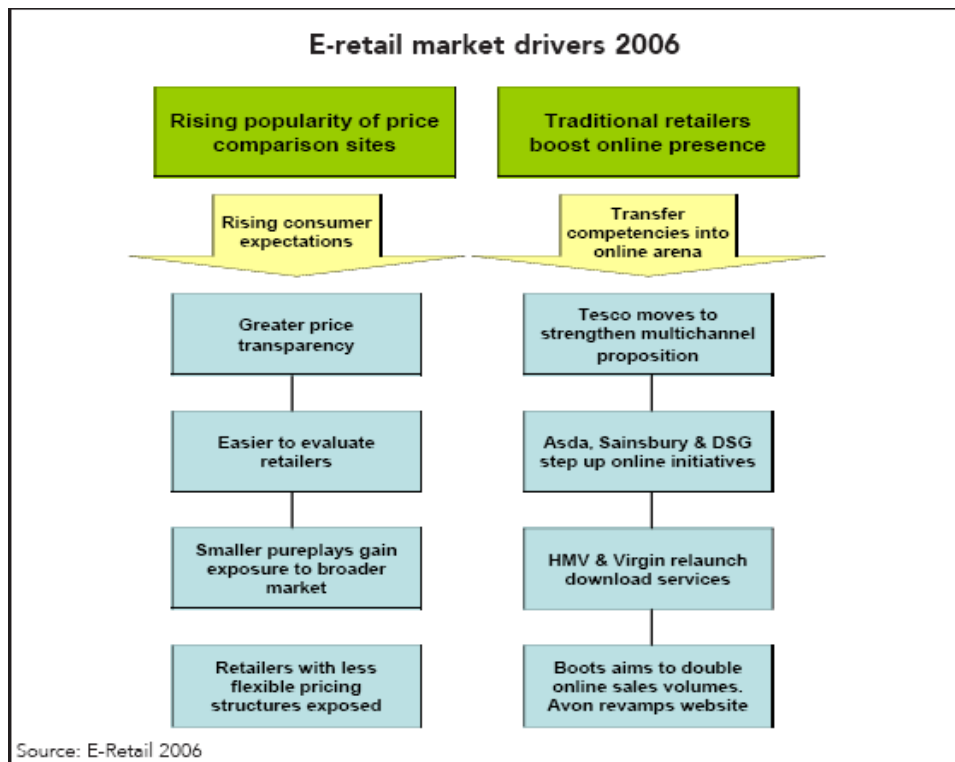
<b>Web Site Characteristics that US Adult Internet Users Find Annoying When They Visit Web Sites to Obtain Information about Products and Services that They Are Interested in Purchasing, July 2005 (as a % of respondents)</b>					
	<b>Extremely annoying</b>	<b>Some-what annoying</b>	<b>Neither annoying nor unannoying</b>	<b>Some-what annoying</b>	<b>Not at all annoying</b>
Pop-up ads	83.8%	9.2%	5.1%	1.3%	0.6%
Requiring the installation of extra software to view the site	72.0%	16.8%	8.2%	1.8%	1.2%
Dead links	66.2%	19.8%	11.1%	2.2%	0.7%
Requirement to register and log-on before viewing the Web site	61.1%	21.9%	12.1%	2.9%	2.0%
Slow-loading pages	57.3%	25.8%	13.6%	2.5%	0.9%
Content that is out of date	56.5%	25.9%	14.5%	2.5%	0.6%
Confusing navigation (hard to find pages, too many clicks)	55.4%	29.0%	13.1%	2.0%	0.6%
Ineffective site search tool	52.4%	27.8%	15.7%	2.8%	1.2%
No contact information available (Web form only)	50.2%	24.9%	19.0%	4.4%	1.5%
Inability to use the browser's "back" button	49.0%	27.0%	17.2%	4.6%	2.2%
Overdone sites (unnecessary splash/flash screens or animation)	41.5%	27.9%	22.4%	5.7%	2.5%
Text that moves	33.5%	26.4%	27.4%	9.2%	3.5%
Music or other audio that plays automatically	30.1%	22.4%	28.3%	12.5%	6.7%
Poor appearance (colors, fonts, format)	28.3%	26.6%	33.2%	8.9%	2.9%
Opening a new window for a link	18.6%	19.8%	33.0%	16.0%	12.6%
<i>Note: n=2,432</i>					
<i>Source: Taylor Nelson Sofres (TNS) commissioned by Hostway, July 2005</i>					
066153 ©2005 eMarketer, Inc. <a href="http://www.eMarketer.com">www.eMarketer.com</a>					

Adapted from [www.emarketing.com](http://www.emarketing.com).

## 7.0 E-retailing facts and figures

According to the "e-Retail 2006" report released by Verdict, one of the most reputable survey and statistics firms in UK, the online market is the fastest growing sector of UK retail, accounting for almost half the cash growth in retail spending. Whilst its share of the retail market is still small, e-retail's meteoric growth is sending shockwaves throughout the wider business sector. The report observes that consumer confidence and expectations are increasing: one in four consumers now purchase goods over the internet, meaning that continual development of an online presence is essential for maintaining and increasing market share. The major drivers of e-retail identified in the Verdict report are listed in figure 2.





**Figure 2.** Verdict (2006: 1).

Among its principal findings, the report concluded that:

*7.1 E-retail is redefining how people select their retailers.* This trend means that e-retail enables consumers to choose products that precisely meet their requirements and to shop at a time convenient to them. This makes online retailers formidable competitors to high street retailers.

*7.2 One in four consumers now purchases goods over the Internet.* As their confidence in their online operations gains momentum, traditional retailers must take steps to establish or develop e-business operations in order to cater to heightened consumer expectations.

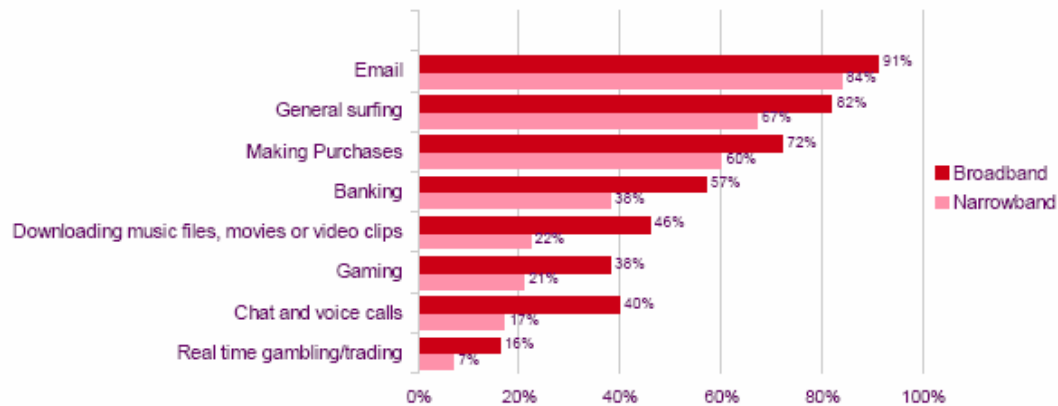
*7.3 Traditional retailers need to leverage the strengths in their offline operations and transfer them to the online arena.* Further development of an online presence is crucial to maintain competitive advantage in the marketplace.

*7.4 Price comparison sites are becoming an increasingly important tool for consumers seeking to compare online retailers.* To possess gain further market share, these sites must develop appropriate attributes.

The argument that businesses should include e-retailing as part of their management strategy is supported by recent developments in the area of internet usage and online shopping. The Internet Statistics Compendium (updated in November 2006 and published by the E-Consultancy) delineates the global reach and penetration of internet services as follows (see figure 3):

- Total internet users in Europe and in the world:  
Europe: 290,121,957  
World: 1,018,057,389
- Total internet users by country and share of world users:

- US: 202,888,307 (21.6%)
  - China: 103,000,000 (11%)
  - Japan: 78,050,000 (8.3%)
  - Germany: 47,127,725 (5.0%)
  - India: 39,200,000 (4.2)
  - UK: 35,807,929 (3.8%)
- Predicted number internet users worldwide: 1.07 billion by 2005, 1.21 billion by 2006, and 1.35 billion by 2007
  - Approximate percentage of UK broadband users who make online purchases: 72%



**Figure 3.** UK online usage statistics, e-Report 2006

## 8.0 Data collection and analysis

This analysis is divided into two sections. The first section is concerned with results of the online survey designed to obtain opinions on a number of issues relevant to this research. Factors assessed include the age groups of the respondents, access to the internet, level of involvement in online product/service search and/or purchase, category of product/service purchased, applicable retail options, shopper's category, e-retailing acceptance or non-acceptance, and respondents' opinions of various retail options.

The second section analyses the survey results in relation to the research objectives. The total number of participants for this survey was 43. Questions 1 and 2, which contain personal details of the participants, have been withheld for ethical reasons.

### 8.1 Survey results

#### 8.1.1 Age group of participants

Out of the 44 participants in this survey, nine (20%) were in the 16–21 years category; 19 (42.2%) were in the 22–30 years category; 12 (26.7%) were in the 31–45 years category; five (11.1%) were in the 46–60 years category; and there were no respondents from the 60 years and above category.

#### 8.1.2 Level of access to the internet

As the vehicle for e-retailing implementation, there is a foundational need to determine the participants' level of access to internet facilities. An overwhelming majority of the respondents—88.4%—indicated that they had constant access to the internet; 7% indicated that they did not always have constant access to the internet; and 4.7% indicated that they had no internet access.

### *8.1.3 Use of internet search facilities to locate products/services*

Participants were questioned about their level of reliance on internet search facilities for locating products or services, as well as the methods they used to carry out their searches. The results show that 77.3% of the participants used online search facilities to locate products or services; 13.6% indicated that they do not; and 9.0% gave other responses, such as "sometimes" and "not always".

When questioned about their preferred ways to carry out internet searches, 31% of the participants indicated that they used search engines (e.g., Google.co.uk, Live.com, AskJeeves.com); 24.1%, used price comparison websites (e.g., Pricerunner.co.uk, moneysupermarket.com); 16.4% used hyperlinks from other websites; 21.6% went directly to the company's website; and 6.9% utilised intranets or other means that they could not remember. A majority of the respondents indicated that they used more than one search method.

### *8.1.4 Participation in e-retailing*

When questioned, 75% of the respondents indicated that they had participated in e-retailing and had made an online purchase; 24.5% had never purchased online, but may have used internet search facilities; 4.5% could not remember if they had made an online purchase in the past.

A second set of questions revealed that electronics and electrical items (e.g., phones and computers) accounted for 16.2% of the respondents' e-retail purchases; consumables, 5.9%; clothing, 15.4%; literature, 9.6%; financial products (e.g., insurance, loans, and credit cards), 13.2%; recreation packages (e.g., holidays, hotel bookings, and tickets), and 12.5%; "do-it-yourself" (DIY) products, 11.8%. The remainder of the purchases—15.4%—were comprised of the following categories: shoes and accessories, cosmetics, an exercise kit, antiques, collector items, a motor scooter on eBay, furniture and ornaments, pleasure toys, academic materials, "none," "cannot remember," and "virtually everything".

### *8.1.5 Frequency of e-retail purchases*

The frequency with which the respondents made e-retail purchases is as follows: 5–10 times a month, 22.7%; 31.8%, 1–4 times a month; 11.4%, 1–5 times in the last six months; 9.1%, 1–5 times in last year; and 25% had never made an online purchase.

### *8.1.6 Participants' opinion of e-retailing adoption by their high street shops*

Participants were asked if they made frequent purchases from a high street store; if that high street shop has an e-retailing component; if they had any advice for high street brands regarding e-retailing; and finally about their views of e-retailing as a management strategy. The results show that 79.5% purchased regularly from a high street shop; 15.9% did not; and 4.5% chose "other".

When asked whether their favoured high street shops have e-retailing capabilities, 29.5% of the respondents said yes; 36.4% said no; and 34.1% indicated that they were not certain.

Respondents were also asked whether they thought that their favoured high street brands should implement an e-retailing strategy for the convenience of their customers. The results—which are quite important for this survey—showed that 75% said yes; 11.3% said it would be okay; 6.8% were indifferent; and 6.8% said definitely no.

Finally, when queried about their views of e-retailing as a business strategy, 59.1% saw it as a marvellous innovation; 22.7% regarded it as okay; 9.1% were indifferent; and 9.1% saw it as a nuisance.

### *8.1.7 Shoppers' categories*

Our survey found that 22.7% of the respondents indicated that they are mainly online shoppers; 45.5%—the majority—indicated that they shop both on the high street and online; 15.9% said that they browse for products and services online, but shop on the high street; 13.6% shop exclusively on the high street; and 2.3% chose "other".

### *8.1.8 Quantitative and qualitative assessment of respondents' views of e-retailing adoption*

Our survey found that 59.1% of respondents were strongly in favour of e-retailing adoption; 27.3% were moderately in favour; 6.8% were against it; 4.5% were strongly against.

Finally, our qualitative analysis of the responses indicates a wide range of opinions about e-retailing strategy, ranging from the resoundingly affirmative to the strongly negative. The reasons given varied widely, depending upon personal experience and direct and indirect observations. A significant majority of the respondents supported the adoption of e-retailing, but also suggested a number of measures that they would like to see implemented to enhance the strategy and make it a safe and more attractive retail option for consumers.

## *8.2 Survey analysis*

Our survey was designed to determine the level of customer acceptance of the e-retailing concept, and to serve as an advisory framework for organisations and business considering its adoption or non-adoption.

In the opinion of many researchers, it is not a matter of "if" non-adopters should implement an e-retailing strategy, but rather "when" they should do so. Adoption is a matter of urgency, they argue, due to the paradigm shift from high street retailing to online retailing, which is rapidly redefining the retail landscape (Verdict, 2006). Our study is likewise premised on the conviction that e-retailing strategy adoption is of paramount importance to organisations hoping to gain competitive advantage or remain relevant in the retail industry. We also argue that the e-retailing strategy should be implemented alongside other existing retail strategies, thereby affording organisations various ways to cash in.

### *8.2.1 Significance of age groups*

A largest share of the respondents to our survey (42.2%) fell into the 22–30 years cohort. The second largest group (26.7%) was comprised of those aged 31–45 years, followed by those in the 46–60 years category (11.1%). There were no respondents in the 60 years and above cohort, due mainly to the researchers' limited number of contacts in that age group. Our survey indicated that across all age group, respondents were widely involved in e-retail-related activities. A vast majority indicated that they have had some form of internet-enabled retail experience, ranging from browsing to online shopping as a main form of retail experience.

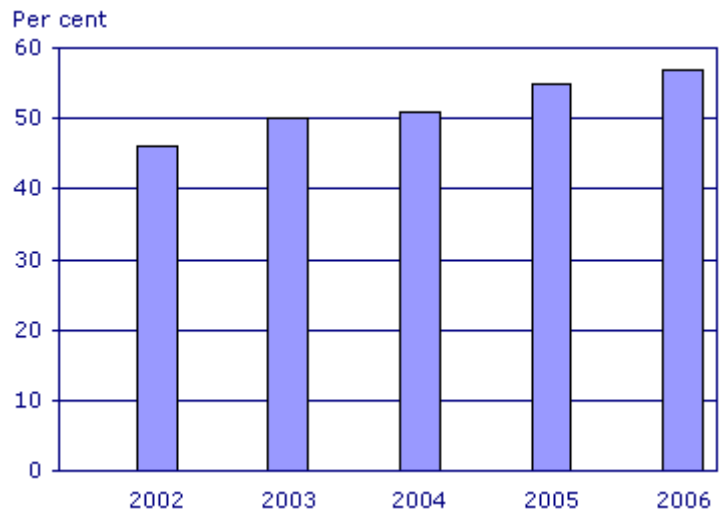
A more in-depth survey was carried out by Verdict, a noted retail research company. According to its "e-Retail 2006" report, the fastest growing age cohort in the UK to embrace the e-retailing concept is comprised of those over age 55: "In 2005, it is the over 55s that offer online retailers the most potential. Unsurprisingly this age group has been the slowest to embrace the Internet, but is now realising its potential. The number of shoppers in this age group rose by 88.5% in the last year to 2.7m—equalling the number of shoppers in the 15–24 age groups" (Verdict, 2006: 3).

This underscores the fact that an increasing number of people are becoming convinced of the benefits of online shopping and are rapidly adopting it. Indeed, whilst there are some people who remain opposed, indifferent, or unaware of the e-retailing concept, the majority now recognise that it is here to stay and are joining the bandwagon.

### *8.2.2 Access to the internet*

A crucial determinant of e-retailing usage is the presence and consistency of internet access for the target user. In our survey, over 80% of the respondents indicated that they had

constant access to the internet. Most had internet access at home, some at school, and others at their places of work. The majority had access to the internet at some point daily (see Figure 4).



**Figure 4.** UK households with access to the internet (Adapted from National Statistics Omnibus Survey, January–April 2006).

### 8.2.3 Internet search facilities enabling e-retailing

E-retailing strategy is facilitated by well-developed internet search facilities. 70% of our survey respondents indicated that they had used such tools to locate and compare products and services (see Figure 4 above). Internet search tools are becoming increasingly popular because they enable buyers to browse amongst available items without visiting an offline shop; compare prices; locate the closest to retail outlet for an offline purchase; and take advantage of special online offers or discounts.

### 8.2.4 Buyers' acceptance of and participation in e-retailing

Backed up by the findings of the "e-Retail 2006" report (Verdict, 2006), this section analyses the level of buyers' acceptance of and participation in e-retailing (see Figure 5 above). That the online retail market is growing rapidly is a fact that organisations and retailers cannot afford to overlook.

An important aim of the survey was to determine which part of the retail industry the respondents patronised most frequently when shopping online (see Figure 6 above). Our results suggest that more and more consumers are finding online shopping to be an attractive avenue for making purchases.

The findings of Verdict's "e-retail 2006" report show that the e-retail market is currently the fastest growing retail sector in the UK (see Figure 5). In 2005, it accounted for almost half the cash growth in retail spending. Indeed, one in four consumers now purchases goods over the internet, meaning that continual development of an online presence is essential if retailers are to maintain and increase market share.



**Figure 5.** Number of online shoppers, 1998–2005 (Adapted from “e-Retail 2006” (Verdict, 2006)).

“The number of online shoppers, here defined as those who had bought ‘retail goods’ over the Internet in the 12 months to October 2005 (that is excluding the likes of air fares, event tickets and insurance) rose by 24.5% to 14.6m [in the UK alone]. The rise was the result both of increased broadband access and increased customer recognition of the benefits of shopping online, in terms of convenience of having goods delivered to the home and access to low price retailers” (Verdict, 2006).

#### 8.2.5 E-retailing versus the high street

The frequency with which our survey respondents made online purchases is outlined in 8.1.5 above. The results indicate that a significant percentage (22.5%) are frequent online shoppers. They also showed that a majority of the respondents utilised both conventional high street shopping and online shopping. This suggests that there is a considerable degree of mutual dependence between offline and online retail channels. Whilst both have their respective costs and benefits, high street retailers need to craft a sound multichannel strategy if they are to successfully broaden their reach beyond the confines of their offline estates (Verdict, 2006).

#### 8.2.6 How consumers view e-retailing and their advice to organisations considering adopting it as a strategy

Our survey respondents gave various reasons for being pro- or anti-e-retailing. Their opinions can be summarised as follows:

##### *Pro:*

- It is a great idea and an important innovation.
- It is easier to purchase an item online and then pick it up at the shop.
- It is much more convenient than shopping on the high street. And it really saves money.
- Facilitates the purchase of goods from places to which that the buyer is unable to go.
- Definitely good for business and very convenient for customers.
- It is good for the cyber-shopper, but a high street presence should still be available for those who need it.
- It is really good to have my shopping come to me rather than the other way round.
- A most essential retail strategy; guaranteed to make high street shopping obsolete in the near future.

- Provides resources that are beyond capabilities of the high street stores.
- It is a welcome, convenient, and less stressful alternative to high street shopping.
- The best deals are online and you can view the full specifications before purchasing.
- E-retailing strips away the power of monopolies like Tesco and ASDA. It provides a level playing field, where virtually anyone can sell anything and the best bargains can be found.
- Adopt it or go out of business.

*Anti:*

- As with other ill-thought-out concepts, this "online shopping thing" will soon go away and take a lot of people down with it. What is wrong with going to the shops to buy your stuff? This is just plain lazy!
- As far as I know, it's just another marketing thing.
- All one needs to do is read about the rapid rise of identity theft and internet payment fraud. This is one of the worst marketing strategies a retailer could come up with. It is just not worth the hassle. Better safe than sorry.
- I prefer to buy my stuff from shops where I can see and feel what I am buying and feel secure about my purchases. I guess I am too scared to give out my details on the internet.
- What is the big deal, really? This is an American thing; they want everything yesterday. What about the tradition of going shopping and the attendant social benefits that make the high street such an important part of our culture? I think this "online shopping" thing is grossly overblown.

Our survey respondents offered a number of recommendations to e-retailers, including:

- Provide reasonable after-sales service to customers
- Offer better security against fraudulent activities by sellers and buyers
- Retain the high street as an option for those who prefer to search the bargain tables
- Eliminate or minimise drastically postage costs and increase the speed of delivery

### **8.3 Conclusions and recommendations**

Our survey took the form of an online questionnaire that was distributed to respondents living in the UK. Its aim was to determine consumers' opinions about e-retailing and e-retailing adoption. The results suggest that e-retailing has an important place in the retail strategies of businesses. It is hoped that organisations will find this research useful as a frame of reference for designing or redesigning their retail strategies. In what follows, we offer some recommendations along these lines.

#### *8.3.1 Generate an e-retailing programme*

The first step is to determine which market is being targeted and to generate an appropriate marketing programme. In the past, retailers have based their buying decisions and marketing programmes on their knowledge of their customers' purchasing preferences. With e-retailing, customers have to be assessed and marketed to on an individual basis (Anderson et al., 2003). For example, Dell.com allows prospective customers to design and configure their computers themselves, while also offering them the option of prefabricated systems available at a wide range of prices.

We recommend that e-retailers implement a system for data capture and construction of user profiles, which would allow them to automatically generate a marketing programme tailored to each individual visiting their website.

#### *8.3.2 Implement cost-reducing marketing strategies*

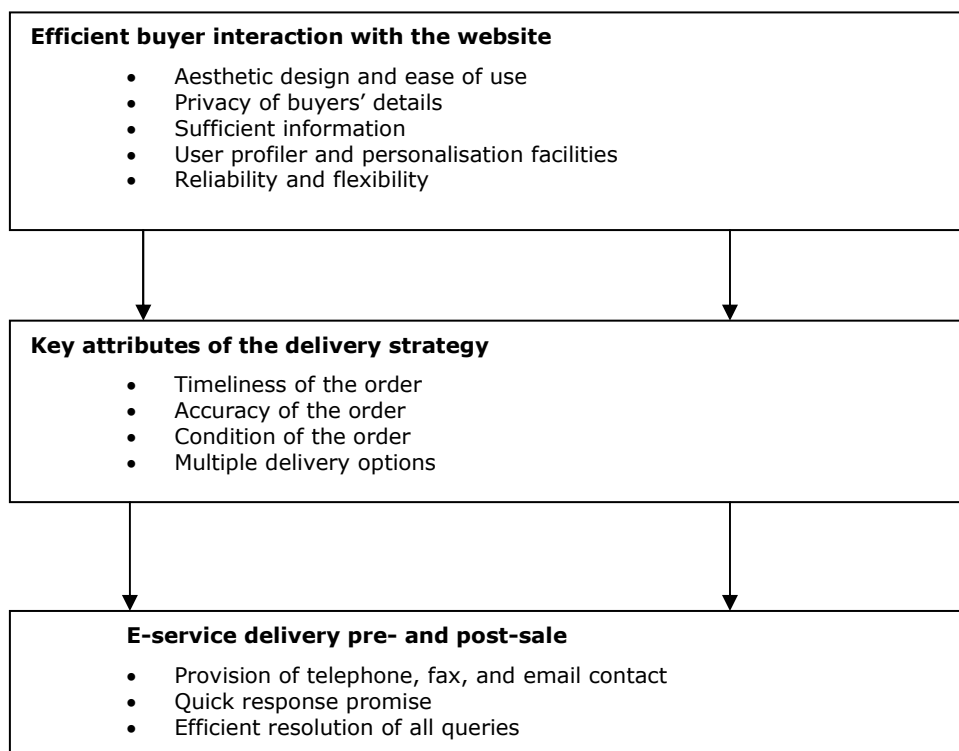
Cost-reduction strategies should be considered hand-in-hand with e-retailing adoption. The organisation will have to decide upfront what its retail model should be. It will need to

implement strategies such as disintermediation, whereby intermediaries such as regional distributors and high street retailers are eliminated from the chain of exchange between producers and consumers. In order to expose their websites to more eyes, a process of re-intermediation might be needed, whereby e-retailers pay content aggregators like Yahoo.com to steer customers to their websites (Bakos, 2001). Most of the respondents to our survey indicated that they locate products and services by using price comparison websites, links from associated websites, or going directly to the website. This suggests that strategies such as disintermediation and re-intermediation are a must for e-retailers.

### 8.3.3 Design an optimum delivery channel strategy

In online retailing, the final stage of the consumer purchase process—the point of delivery—is the most important, yet the most undeveloped, of retailers’ internet operations. Retailers need to ensure that they perfect this vital stage in the process. Failing to do so can result in adverse customer perceptions of online as a reliable and efficient purchasing channel; it can also impact their perceptions of the e-retailers’ integrity (Verdict, 2006). There is thus a need to offer more delivery options for consumers and to liaise closely with logistics operators. This delivery strategy needs to be prioritised to ensure long-term credibility.

Figure 6 outlines a number of recommended strategies aimed at assuring the success of e-retailing operations.



**Figure 6.** Points for strategy development

## 9.0 Future research

This research was designed to consider buyer opinions of e-retailing strategy adoption by organisations. There is need for further research to address questions such as, what is the cost-to-profit ratio of e-retailing strategy adoption? How will mainline high street brands prevent channel conflict if e-retailing strategy is adopted? And what are the long-term prospects for the e-retailing strategy?



## References

- Accenture (2001) 'European logistics companies adopting "wait and see" attitude to web technology risk being left behind. Available at: [http://accenture.tekgroup.com/article\\_display.cfm?article\\_id=3781](http://accenture.tekgroup.com/article_display.cfm?article_id=3781).
- Anderson, W. P., Chatterjee, L., and Lakshmanan, T.R. (2003) 'E-commerce, transportation and economic geography, *Growth and Change* 34 (4), pp. 415–432.
- Annenberg Public Policy Center (2003) 'Most Americans do not understand how websites use information about them.' Available at: <http://www.upenn.edu/pennnews/researchatpenn/article.php?686&soc>.
- Barsh, J., Crawford, B., and Grosso, C. (2000) 'How e-tailing can rise from the ashes,' *McKinsey Quarterly* 3, pp. 98–109.
- Berry, J. (2005) 'Quantitative methods in education research.' Centre for Teaching Mathematics, University of Plymouth. Available at: <http://www.edu.plymouth.ac.uk/RESINED/Quantitative/quanthme.htm>
- Bryman, A. (2001) *Social Research Methods*. Oxford: Oxford University Press.
- Bakos, Y. (2001) 'The emerging landscape for retail e-commerce,' *Journal of Economic Perspectives* 15 (1), pp. 69–80.
- Cagliano, R., Caniato, F., and Spina, G. (2003) 'E-business strategy: how companies are shaping their supply chain through the Internet,' *International Journal of Operations and Production Management* 23 (10), pp. 1142–1163.
- Cassidy, C.M., Chae, B. (2006) 'Consumer information use and misuse in electronic business: an alternative to privacy regulation,' *Information Systems Management* 23 (3), pp. 75–88.
- Chaffey, D. (2004) *E-Business and E-Commerce Management, 2<sup>nd</sup> Edition*. Essex: Prentice Hall.
- Chaffey, D., Mayer, R., Johnston, K., and Ellis-Chadwick, F. (2000) *Internet Marketing: Strategy, Implementation and Practice*. Harlow: Prentice Hall.
- Chen, S. (2005) *Strategic Management of e-Business*. West Sussex: John Wiley & Sons.
- Collier, E.J. and Bienstock, C.C. (2006) 'How do customers judge quality in an e-tailer?,' *MIT Sloan Management Review* 48 (1), pp. 35–40.
- Côté, L., Vézina, M., and Sabourin, V. (2005) 'The strategic management process in e-business,' *Ivey Business Journal* May/June, pp. 1–8. Available at: [http://www.iveybusinessjournal.com/view\\_article.asp?intArticle\\_ID=561](http://www.iveybusinessjournal.com/view_article.asp?intArticle_ID=561).
- Croom, S.R. (2005) 'The impact of e-business on supply chain management: an empirical study of key developments,' *International Journal of Operations and Production Management* 25 (1), pp. 55–74.
- Davis, J. (2005) 'Types of Research Methods.' Available at: <http://www.naropa.edu/faculty/johndavis/prm2/types1.html>.
- Dawson, J. (2000) 'Retailing at century end: some challenges for management and research,' *The International Review of Retail, Distribution and Consumer Research* 10 (1), pp. 119–148.
- Denzin, N.K. and Lincoln, Y.S., eds. (1994) *Handbook of Qualitative Research*. London: Sage Publications.
- Dixons (2006) 'Dixons—part of the UK's biggest electrical retailer.' Available at: <http://www.dixons.co.uk/martprd/editorial/DSG%20international%20PLC>.
- E-Consultancy (2006) 'Internet statistics compendium.' Available at: <http://www.e-consultancy.com/publications/internet-stats-compendium>.
- Edgecliffe-Johnson, A. (2000) 'E-revolution shelved,' *The Financial Times*. Available at: <http://specials.ft.com/In/specials/spe05a.htm>.
- Ernst and Young (2001) 'Global online retailing.' Available at: [http://www.ey.com/global/Content.nsf/South\\_Africa/16\\_Jan\\_01\\_On-line\\_Retail](http://www.ey.com/global/Content.nsf/South_Africa/16_Jan_01_On-line_Retail).

- Exley, L., Bidwell, J., Murphy, J. and Brewer, J. (2006) 'Is the future of retail online or can the high street be revived?' *Marketing* 4 October, p. 24.
- Evans, P. and Wurster, T.S. (2000) *Blown to Bits. How the New Economics of Information Transforms Strategy*. Boston: Harvard Business Press.
- Frohlich, M.T. and Dixon, J.R. (2001) 'A taxonomy of manufacturing strategies revisited,' *Journal of Operations Management* 19 (1), pp. 541–558.
- Foresight (2000) 'Clicks and mortar: the new store formats,' London: Department of Trade and Industry.
- Hanford, D.J. (2005) 'Broadband boosts e-retailing,' *Wall Street Journal* 8 June, p. 5.
- Jiang, P., and Talaga, J. (2006) 'Building a customer base in the electronic marketplace: an empirical exploration of the e-tailing industry,' *The Journal of Services Marketing* 20 (7), p. 429.
- Joia, L.A. and Sanz, P. (2005) 'The hidden value of sporadic customers in e-retailing: an empirical investigation,' *Internet Research* 15 (1), pp. 7–14.
- Jones, P., Clarke-Hill, C., and Hillier, D. (2002) '(R)etailing in the UK,' *Marketing Intelligence and Planning* 20 (4/5), pp. 229–304.
- Johnston, R. and Lawrence, P.R. (1988) 'Beyond vertical integration: the rise of value adding partnership,' *Harvard Business Review* 66 (4), pp. 94–101.
- Kaplan, B. and Maxwell, J.A. (1994) 'Qualitative research methods for evaluating computer information systems,' in Anderson, J.G. et al., eds. *Evaluating Health Care Information Systems: Methods and Applications*. Thousand Oaks, CA: Sage Publications.
- Kapuscinski, R., Zhang, R.Q., Carbonneau, P., Moore, R., and Reeves, R. (2004) 'Inventory decisions in Dell's supply chain interfaces,' *Linthicum*. 34 (3), pp. 191–206.
- Kosynski, B.R. and McFarlan, F.W. (1990) 'Information partnerships: shared data, shared scale,' *Harvard Business Review* 68 (5), pp. 114–120.
- Laudon, M. and Laudon, J. (2003) 'The digital firm: e-business and e-commerce,' in *Management Information Systems: Managing the Digital Firm, 8<sup>th</sup> International Edition*. Englewood Cliffs, NJ: Prentice Hall.
- Laudon, M. and Laudon, J. (2005) 'Essentials of management information systems,' in *Managing the Digital Firm*. Englewood Cliffs, NJ: Prentice Hall.
- Leonard, L.N.K. and Cronan, T.P. (2003) 'Website retailing: electronic supply chain replenishment,' *Journal of End User Computing* 15 (3), p. 45.
- Lee, H.L. and Whang, S. (2001) 'Supply chain integration over the Internet.' Available at: <http://www.commerce.net/research/ebusiness-strategies/2kl/2kl-2kl-r.html>
- Malone, T., Yates, J. and Benjamin, R. (1987) 'Electronic markets and electronic hierarchies: effects of information technology on market structure and corporate strategies,' *Communications of the ACM* 30 (6), pp. 484–497.
- Maltz, A., Rabinovich, E., and Sinha, R. (2004) 'Logistics: the key to e-retail success,' *Supply Chain Management Review* 8 (3), pp. 56–64.
- Merrilees, B. and Miller, D. (2005) 'Emotional brand associations: a new KPI for e-retailers,' *International Journal of Internet Marketing and Advertising* 2 (3), p. 206.
- Miles, M.B. and Huberman, A.M. (1994) *Qualitative Data Analysis: An Expanded Source Book, 2<sup>nd</sup> Edition*. Thousand Oaks, CA: Sage Publications.
- Myers, M.D. (1997) 'Qualitative research in information systems,' *MIS Quarterly* 21 (2), pp 241–242.
- National Consumer Council (2000) *E-commerce and Consumer Protection: Consumers' Real Needs in a Virtual World*. London: National Consumer Council.
- Office for National Statistics (2005) '2005 e-commerce survey of business. Available at: <http://www.statistics.gov.uk/pdfdir/ecom1006.pdf>.
- Office for National Statistics (2009). 'Internet access.' Available at: <http://www.statistics.gov.uk/cci/nugget.asp?id=8>.

- Newbery, M. (2004) 'Trends in online apparel retailing, forecasts to 2010—Management briefing: How far can online go?,' *Just-Style* 11 (1), pp. 22–25.
- Orlinkwski, W.J. and Baroudi, J.J. (1991) 'Studying information technology in organisations: research approaches and assumptions,' *Information Systems Research* 2, pp. 1–28.
- Pateli, A.G. and Giaglis, A. (2004) 'A research framework for analysing eBusiness models,' *European Journal of Information Systems* 13 (4), p. 302.
- Peng, Y., Trappey, C.V., and Liu, N. (2005) 'Internet and e-commerce adoption by the Taiwan semiconductor industry,' *Industrial Management and Data Systems* 105 (3/4), pp. 476–491.
- Poulsen, K. (2000) 'Survey: security fears are global. *Security Focus*. Available at: <http://www.securityfocus.com/news/106>.
- PR Newswire. (2005). 'Survey captures business impact of banks, online businesses leaving consumers unprotected online.' Available at: <http://www.lexisnexis.com/academic/universe/default.asp>.
- Robson, C. (1993) *Real World Research: A Resource for Social Scientists and Practitioners-Researchers*. Oxford: Oxford University Press.
- Schappell, D. (2000) 'E-tailing.' Available at: [http://searchcio.techtarget.com/sDefinition/0,,sid182\\_gci212079,00.html](http://searchcio.techtarget.com/sDefinition/0,,sid182_gci212079,00.html).
- Shafaghi, M. (2006) 'The economics of e-business.' Lesson 1–12 lecture notes, University of Bolton, UK.
- Smith, P.R. and Chaffey, D. (2005) *eMarketing eXcellence: At the heart of eBusiness, 2<sup>nd</sup> Edition*. Oxford: Butterworth Heinemann.
- Symons, J. (2005) 'Forrester Research US eCommerce forecast: online retail sales to reach \$329 billion by 2010.' Cambridge: Forrester Research.
- Trabold, L.M., Heim, G.R., and Field, J.M. (2006) 'Comparing e-service performance across industry sectors: drivers of overall satisfaction in online retailing,' *International Journal of Retail and Distribution Management* 34 (4/5), pp. 240–258.
- Verdict (2006) 'e-Retail 2006: Online retailers re-mould the landscape.' Available at: <http://www.verdict.co.uk/Marketing/ERET06mktg.pdf>.
- Williamson, O.E. (1979) 'Transaction cost economics: the governance of contractual relations,' *The Journal of Law and Economics* 22 (2) pp. 232–262.
- Zhuang, Y. (2005) Does electronic business create value for firms? an organizational innovation perspective,' *Journal of Electronic Commerce Research* 6 (2), pp. 146–156.
- Zwass, V. (1997) *Foundations of Information Systems*. New York: McGraw-Hill.

# An investigation into the prevalence of *Toxoplasma gondii* infection in *Apodemus sylvaticus* in North Yorkshire

**Nicole Dodd, Denise Thomasson, Elizabeth Wright, Jacqueline Hughes, and Geoff**

**Hide**, Centre for Parasitology and Disease, School of Environment and Life Sciences, University of Salford

## Abstract

The coccidian parasite *Toxoplasma gondii* is ubiquitous and affects many mammals including humans. Wild and domestic felids are the definitive host, however *Toxoplasma* can infect all warm-blooded animals. Toxoplasmosis is a significant cause of abortion, stillbirth and disease in both humans and livestock. Three routes of transmission occur; ingestion of tissue cysts in the intermediate hosts, ingestion of oocysts shed by the definitive host into the environment and the transplacental crossing of tachyzoites from mother to foetus. The relative importance of each of these transmission routes in the epidemiology of the parasite has been long debated, however it is widely accepted that the definitive host is the main source of infection. The wood mouse (*Apodemus sylvaticus*) was sampled as part of a continuous study of parasites in the area over a 10 year period from North Yorkshire in an area free from cats. The presence of *Toxoplasma* was detected using PCR amplification of the SAG-1 gene. Over the 10 year study 81/206 mice tested positive for *Toxoplasma gondii*; the mean prevalence was 39.3% ( $\pm$  6.2%). As the area is free from cats, this high level of prevalence suggests that other mechanisms of parasite transmission are likely to be important in this population of *Apodemus sylvaticus*.

## Introduction

The obligate intracellular parasite *Toxoplasma gondii* is ubiquitous and uses an unusually diverse range of hosts. It is a member of the phylum Apicomplexa characterised by possessing an 'apical complex' located in the anterior end of the organism used for host cell penetration. The phylum is an ancient group comprising of around 5000 parasitic species including significant pathogens of humans and animals. These include the causative agent of malaria *Plasmodium spp* and the common poultry disease coccidiosis caused by *Eimeria spp*. Both wild and domestic felids are the definitive host where the parasite completes its sexual cycle. However *Toxoplasma* can potentially infect all warm-blooded animals.

The parasite is of medical and veterinary importance as it can cause congenital disease and abortion in intermediate hosts (Dubey and Jones, 2008). In humans infection is usually afebrile and self-limiting although immunocompromised hosts can suffer lymphadenopathy or more seriously encephalitis. Toxoplasmic encephalitis is the most common presentation of the disease in AIDS patients and is usually due to the reactivation of latent tissue cysts (Howe and Sibley, 1995).

Three routes of transmission occur: the horizontal transmission of infective oocysts shed by the cat, the ingestion of tachyzoites from tissue cysts and vertically through the transplacental crossing of tachyzoites from mother to foetus (Rosai, 2004). The relative importance of each of these routes of transmission is still unclear. There are many strains of *Toxoplasma* however the vast majority (>95%) fall into three distinct lineages (reviewed in Sibley, 2003), of which the genotypes within each lineage differ by 1% or less (Su et al., 2003). It is unusual that, despite the ubiquity of the parasite, the population structure of *T. gondii* is extremely clonal and highly conserved. This points towards a relatively recent common ancestor followed by predominantly asexual propagation (Dubey and Jones, 2008; Su et al., 2003).

Cats are thought to be the major source of infection although only 1% of cats are estimated to be shedding viable oocysts at any one time (Dubey and Beattie, 1988). Infective meat is also considered important although this does not explain the high prevalences found in strict vegetarians (Hall et al., 1999). Vertical transmission is traditionally considered less important epidemiologically, however it has been observed at high levels in natural populations and experimental infections in numerous species (Duncanson et al., 2001; Flori et al., 2002; Morley et al., 2005; Morley et al., 2008; Roberts et al., 1994; Williams et al., 2005). Indeed, high levels of vertical transmission of *Toxoplasma* has been observed in murine hosts

experimentally infected (Owen and Trees, 1998) and also in natural urban populations (Marshall et al., 2004; Murphy et al., 2008).

The objective of the current study was to investigate the prevalence of *Toxoplasma gondii* in a wild population of wood mouse (*Apodemus sylvaticus*) in North Yorkshire using molecular techniques. Furthermore, we aim to consider the importance of transmission routes that bypass the cat by investigating the prevalence of *Toxoplasma* in an area free of cats.

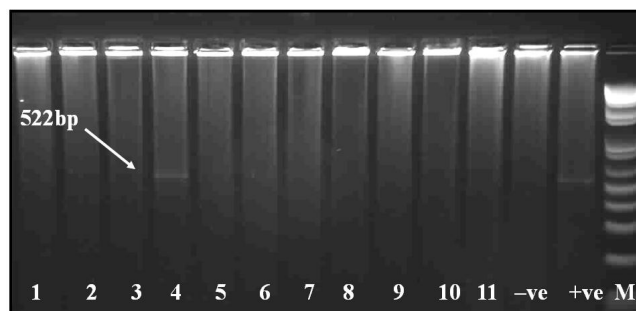
## Materials and methods

The wood mouse (*Apodemus sylvaticus*)  $n = 206$  was sampled from the surrounding area of Malham Tarn, North Yorkshire over a 10 year period as part of a continuous study of parasites in the area. Brains were removed in sterile conditions and stored frozen in 400 $\mu$ l of lysis buffer and then taken back to Salford University. DNA was extracted from brain tissue using phenol/chloroform standard procedure as described by Williams et al. (2005), and the presence of *Toxoplasma* was detected using nested PCR amplification of the Surface Antigen Gene 1 (SAG-1) (Savva et al., 1990). The PCR reaction was conducted as follows: each 25 $\mu$ l of PCR consisted of 2.5 $\mu$ l Bioline (without MgCl<sub>2</sub>), 2.5 $\mu$ l  $\beta$ -mercaptoethanol (50mM), 1 $\mu$ l MgCl<sub>2</sub> (50 mM), 0.25  $\mu$ l of deoxynucleotide triphosphate mix (25 mM each), 2.5  $\mu$ l of each oligonucleotide primer 10 pM/ $\mu$ l, 12.25  $\mu$ l of water, 0.5  $\mu$ l of Taq DNA polymerase (5 units/ $\mu$ l) and 1  $\mu$ l of DNA template. In the first PCR amplification was carried out at 95°C for 5 min, then 40 cycles of (i) 95°C for 40 sec (ii) 63°C for 40 sec and (iii) 72°C for 1 min 10 sec, and lastly 72°C for 10 min with primers 5' TTGCCGCGCCCACTGATG 3' and 5' CGCGACACAAGCTGCGATAG 3'. Following the first round of PCR, 2 $\mu$ l the first round product was used in the second round with the second round primer pairs 5' CGACAGCCGCGGTCATTCTC 3' and 5' GCAACCAGTCAGCGTCGTCC 3', using the same number of cycles and thermal profile as the first round.

Prior to the SAG-1 assay DNA was tested using primers for mammalian tubulin to ensure that the quality of the DNA was suitable for PCR in order to reduce the chances of accepting a false negative. Concentrations of DNA were also increased and decreased to ensure true negativity. All mice collected in 2008 tested positive for tubulin and therefore were able to successfully undergo PCR. In previous years, mice samples that tested negative for tubulin were omitted from the study. PCR products were visualised on 1.5% gels which were trans-illuminated with UV light. Molecular grade water was used in the reactions as a negative control.

## Results

DNA extracted from *Apodemus sylvaticus* was amplified using the SAG1 PCR and analyses by gel electrophoresis. An example is shown in Figure 1. A 522bp band indicates a positive amplification. A total of 206 mice were tested for *Toxoplasma* over a 10 year period. The prevalence infection determined by PCR can be seen from Table 1. An overall prevalence of 39.3% ( $\pm$  6.2%) of *T. gondii* was detected over the 10 years of study. Prevalences ranged between 10.5% (2007) to 69% (2003).



**Figure 1.** SAG-1 PCR gel electrophoresis Malham mice 2008. Lane 1, mouse 280, lane 2, mouse 281, lane 3, mouse 282, lane 4, mouse 283, lane 5, mouse 284, lane 6, mouse 285, lane 7, mouse 286, lane 8, mouse 287, lane 9, mouse 288, lane 10, mouse 289, lane 11, mouse 290, lane 12, mouse 291, lane 13, mouse 292, -ve negative control, +ve positive control, M 1Kb Invitrogen marker. 1  $\mu$ l DNA was used per reaction.

**Table 1.** Prevalence of *T. gondii* infection at Malham Tarn, Yorkshire

Year	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	Mean
Total mice	37	24	19	16	29	8	24	9	19	21	206
Infected	17	4	6	10	20	3	13	4	2	2	81
Prevalence (%)	46	17	32	62.5	69	37.5	54	44	10.5	9.5	39.3

## Discussion

In this study we investigated the prevalence of *Toxoplasma gondii* in a wild population of *Apodemus sylvaticus* in North Yorkshire using a PCR based detection assay. We found 81 of the 206 mice sampled positive for *Toxoplasma*; the mean prevalence over the last 10 years in the area was 39.3%. As the area is free from cats, this suggests that other mechanisms of parasite transmission may be important in this population of *Apodemus sylvaticus*. This raises important questions about the routes of transmission of this parasite. The cat is the only known definitive host however the importance of its role in the epidemiology of *Toxoplasma* has recently been questioned (e.g. Hide et al., 2009). This is largely due to the fact that oocysts are only shed from the feline host for around 14 days as a kitten (Hill and Dubey, 2002). Further to this, certain groups of humans and mammals found to have high prevalence of *Toxoplasma* should, theoretically, have low prevalences due to either their customs or environment restricting them to oocyst exposure. An example of this is the religious group 'Jains' who are strict vegetarians and do not eat root vegetables. Coincidentally the Jain's religious practices follow recommendations similar to those given to pregnant women for prevention of *Toxoplasma* infection (Hall et al., 1999). High levels of *Toxoplasma* have also been found in marine mammals (Mikaelian et al., 2000; Oksanen et al., 1998); approximately 16% of sea otter deaths off the coast of California have been linked to *Toxoplasma* infection (Miller et al., 2004). What is more, the population structure of *Toxoplasma* is clonal and highly conserved, suggesting limited genetic exchange. This would point towards the asexual routes being the significant method of propagation for this parasite, and as the parasite has been detected in strict herbivores, it would seem that vertical transmission may be a grossly underestimated contributor in the parasite's wide distribution and high prevalence.

The high levels of vertical transmission recorded in various mammalian species (Duncanson et al., 2001; Flori et al., 2002; Marshall et al., 2004; Morley et al., 2008; Murphy et al., 2008; Owen and Trees, 1998; Roberts et al., 1994) raises the issue of the possibility that congenital transmission occurs more frequently than currently thought. Further to this, the mechanism by which this happens may actually be the reactivation of latent infections caused by the immunosuppressed state in which is induced by pregnancy, rather than the result of the host being infected for the first time during gestation. The genetic population structure of *Toxoplasma gondii* also suggests limited genetic exchange since a recent event. This would point towards the asexual routes being the significant method of propagation for this parasite. Further research is necessary to elucidate the relative importance of transmission routes in this important pathogen.

## Acknowledgments

We would like to thank the University of Salford for financial support.

## References

- Couper, K.N., Nielsen, H.V., Petersen, E., Roberts, F., Roberts, C.W. and Alexander, J. (2003) DNA vaccination with the immunodominant tachyzoite surface antigen (SAG-1) protects against adult acquired *Toxoplasma gondii* infection but does not prevent maternofetal transmission. *Vaccine*. 21, 2813-2820
- Dubey, J.P. and Beattie, C.P. (1988) *Toxoplasmosis of animals and man*. Boca Raton (FL): CRC Press
- Dubey, J.P. and Jones, J.L. (2008) *Toxoplasma gondii* infection in humans and animals in the United States. *International Journal for Parasitology*. 38, 1257-1278
- Duncanson, P., Terry, R.S., Smith, J.E. and Hide, G. (2001) High levels of congenital transmission of *Toxoplasma gondii* in a commercial sheep flock. *International Journal of Parasitology*. 31, 1699-1703
- Flori, P., Hafid, J. Bourlet, T., Raberin, H., Genin, C. and Tran Manh Sung, R. (2002) Experimental model of congenital toxoplasmosis in guinea-pigs: use of quantitative and qualitative PCR for the study of maternofetal transmission. *Journal of Medical Microbiology*. 51, 871-878
- Hall, S.M., Pandit, A., Golwilar, A., Williams, T.S. (1999) How do Jains get toxoplasma infection? *Lancet*. 354, 486-487
- Hide, G., Morley, E.K., Hughes, J.M., Gerwash, O., Elmahaishi, M.S., Elmahaishi, K.H., Thomasson, D., Wright, E.A., Williams, R.H., Murphy, R.G. and Smith, J.E. (2009) Evidence for high levels of vertical transmission in *Toxoplasma gondii*. *Parasitology*, in press. 139
- Hill, D. and Dubey, J.P. (2002) *Toxoplasma gondii*: transmission, diagnosis and prevention. *Clinical Microbiology and Infectious Disease*. 8, 634-640
- Howe, D.K. and Sibley, D.L. (1995) *Toxoplasma gondii* comprises three clonal lineages: correlation of parasite genotype with human disease. *The Journal of Infectious Disease*. 172, 1561-1566
- Marshall, P.A., Hughes, J.M., Williams, R.H., Smith, J.E., Murphy, R.G. and Hide, G. (2004) Detection of high levels of congenital transmission of *Toxoplasma gondii* in natural urban populations of *Mus domesticus*. *Parasitology*. 128, 39-42
- Mikaelian, I., Boisclair, J., Dubey, J.P., Kennedy, S. and Martineau, D. (2000) Toxoplasmosis in beluga whales (*Delphinapterus leucas*) from the St Lawrence estuary: two case reports and a serological survey. *Journal of Comparative Pathology*. 122, 1, 73-76
- Miller, M.A., Grigg, M.E., Kreuder, C., James, E.R., Melli, A.C., Crosbie, P.R., Jessup, D.A., Boothroyd, J.C., Brownstein, D. and Conrad, P.A. (2004) An unusual genotype of *Toxoplasma gondii* is common in California sea otters (*Enhydra lutris nereis*) and is a cause of mortality. *International Journal for Parasitology*. 34, 3, 275-284
- Morley, E.K., Williams, R.H., Hughes, J.M., Terry, R.S., Duncanson, P., Smith, J.E. and Hide, G. (2005) Significant familial differences in the frequency of abortion and *Toxoplasma gondii* infection within a flock of Charollais sheep. *Parasitology*, 131 181-185
- Morley, E.K., Williams, R.H., Hughes, J.M., Thomasson, D., Terry, R.S., Duncanson, P., Smith, J.E. and Hide, G. (2008) Evidence that primary infection of Charollais sheep with *Toxoplasma gondii* may not prevent foetal infection and abortion in subsequent lambings. *Parasitology*. 135, 169-173
- Murphy, R.G., Williams, R.H., Hughes, J.M., Hide, G., Ford, N.J. and Oldbury, D.J. (2008) The urban house mouse (*Mus domesticus*) as a reservoir of infection for the human parasite *Toxoplasma gondii*: an unrecognised public health issue? *International Journal of Environmental Health Research*. 18, 3, 177-185
- Oksanen, A., Tryland, M., Johnsen, K. and Dubey, J.P. (1998) Serosurvey of *Toxoplasma gondii* in North Atlantic marine mammals by use of agglutination test employing whole tachyzoites and dithiothreitol. *Comparative immunology, Microbiology and Infectious Disease*. 21, 107-114

- Roberts, C.W., Brewer, J.M. and Alexander, J. (1994) congenital toxoplasmosis in the Balb/c mouse: prevention of vertical disease transmission and foetal death by vaccination. *Vaccine*. 12, **15**, 1389-1394
- Rosai, J. (2004) Rosai and Ackerman's Surgical Pathology. *Elsevier*. 1, 9th Edition. London, UK
- Savva, D., Morris, J.C., Johnson, J.D., and Holliman, R.E. (1990) Polymerase chain reaction for the detection of *Toxoplasma gondii*. *Journal of Medical Microbiology*. 32, 25-31
- Sibley, L. D. (2003) Recent origins among ancient parasites. *Veterinary Parasitology*. 115, 185-198
- Su, C., Evans, D., Cole, R.H., Kissinger, J.C., Ajioka, J.W. and Sibley, L.D. (2003) Recent expansion of *Toxoplasma* through enhanced oral transmission. *Science*. 299, 414-416
- Tenter, A.M., Heckeroth, A.R. and Weiss, L.M. (2000) *Toxoplasma gondii*: from animals to humans. *International Journal of Parasitology*. 30, 1217-1258
- Williams, R.H., Morley, E.K., Hughes, J.M., Duncanson, P., Terry, R.S., Smith, J.E., Hide, G., (2005) High levels of congenital transmission of *Toxoplasma gondii* in longitudinal and cross-sectional studies on sheep farms provides evidence of vertical transmission in ovine hosts. *Parasitology*. 130, 301-307



# Lexical, Cultural, and Grammatical Translation Problems Encountered by Senior Palestinian EFL Learners at the Islamic University of Gaza, Palestine

Mohammed El Haj Ahmed, *School of Languages, University of Salford*

## Abstract

The purpose of this empirical study is to investigate, analyze and classify the lexical, cultural and grammatical problems encountered by senior Palestinian EFL university learners in their English-Arabic translations and to find out any gender differences regarding their translation problems. The translation test used in this study consisted of 18 texts covering three types: descriptive, argumentative and instructional. These texts were translated by 28 students chosen randomly from the Department of English at the Islamic University of Gaza. The students also responded to a questionnaire regarding the level of difficulty of each translated text. The analysis of students' translations was based on the linguistic judgments of a group of Arabic specialists including the researcher's assessment, a rater's assessment, the target readership assessment and the work of a professional translator. The results showed that Palestinian EFL student translators committed errors on three levels: lexical, cultural and grammatical. The most frequent errors were grammatical errors, followed by lexical errors and cultural errors. Grammatical errors were mainly attributed to linguistic differences between English and Arabic, as well as students' lack of linguistic competence. Lexical errors were mainly attributed to students' tendency to adopt literal translation and their total dependence on existing bilingual dictionaries. Cultural errors on the other hand resulted from students' unfamiliarity with the source language culture and lack of knowledge of the target culture. The study also revealed that male students committed more errors than female students at lexical and grammatical levels. However, male students committed fewer errors than female students at the cultural level. The results of the questionnaire showed that argumentative texts were regarded by students as the most difficult, followed by instructional and finally descriptive texts.

## 1. Introduction

Translation from one language to another frequently causes errors which result in miscommunication of the original text. Errors often result from language transfer which assumes a one-to-one correspondence of literal meaning between two languages (Wallmach & Kruger, 1999). Language transfer creates negative effects in translation by introducing unnatural sentences which may often be incomprehensible or inadequate to the readers of the target language. This phenomenon, called "translationese" by Nida (1964), is created by a translator's excessive fidelity toward the source language and is now viewed as one of the main problems in translation.

In translating from English into Arabic the translator is dealing with two languages that are unrelated both linguistically and culturally. Dickins et al. (2002: 29) state: "translating involves not two languages, but a transfer from one culture to another". The same view is shared by Mailhac (1995: 1) who maintains that effective communication can be achieved if translation is capable of bridging the gap not only between languages but also between cultures. In translating into Arabic the translator may face linguistic and cultural difficulties that would lead to the distortion of the work being translated or would affect the readers' response in the target language. For example, in Arabic the word *بومة* has many negative connotations and is always seen as a symbol of bad luck, while in English, 'owl' has positive and favourable associations, and is seen as a symbol of wisdom and grace. According to Nida (1964), there is no semantic unit that has exactly the same meaning in two different utterances; there are no complete synonyms within a language; and there is no exact correspondence between related words in different languages. Thus, a word-for-word translation that closely follows the form of the source language is of little or no value to the readers of the target language. For example, to translate the English expressions 'He is as wise as an owl' or 'He is a wise old owl' into Arabic as *هو حكيم كالبيومة* would be unacceptable to the target language reader because of the negative connotations of *بومة* in Arabic.

One means of assessing whether translation problems have been overcome, is by deciding the extent to which target the translated text effectively fulfills its intended role in the target language and culture rather than just evaluating it in terms of an accurate rendering of the source text. In this light, a 'good' translation is one which is acceptable, 'adequate', or 'functionally appropriate' to the native speakers of the target language, as Schaffner (1998) noted. The judgment of qualified target language experts of the translated material is one of the most significant criteria to determine the quality of the translation.

### **1.1. Research questions**

This study aims at identifying, analyzing, and classifying types of lexical, cultural and grammatical problems in Palestinian EFL university students' English-Arabic translation and viewing them in terms of acceptability as judged by a group of qualified experts in Arabic. Specifically, the study addresses the following questions:

1. What types of lexical errors are most frequently committed by senior Palestinian EFL learners?
2. What types of cultural errors are most frequently committed by senior Palestinian EFL learners?
3. What types of grammatical errors are most frequently committed by senior Palestinian EFL learners?
4. Is there any gender difference regarding the translation errors committed by the subjects of this study?

### **1.2. Need for the research**

English is now a major international language and the demand for translation has steadily increased. Moreover, English is the language of instruction at local and international universities, which means that a large number of materials are translated into English as well as other languages including Arabic. However, the lack of qualified translators and the absence of effective and systematic training in translation theory have led to erroneous translations. In this regard, Masoud (1988: 10) is right when she says:

More often than not, new translators dive into translation work thinking that because they speak two languages, they are qualified for the task.

Masoud's statement makes it clear that translators depend on their language experience to practice translation, a view which is also shared by Baker (1992: 4), who states:

Translators need to develop an ability to stand back and reflect on what they do and how they do it. Like doctors and engineers, they have to prove to themselves as well as others that they are in control of what they do; that they do not just translate well because they have a 'flair' for translation, but rather because, like other professionals, they made a conscious effort to understand various aspects of their work (ibid: 4).

In the Department of English, IUG (where the author of this study worked as a lecturer of English for 10 years), the two translation modules which are a major requirement for the degree of BA in English Language and Literature are taught by specialists in EFL and linguistics; none of them has a qualification in translation. Given this situation, the graduates of departments of English will join the profession of translation dependent on their intuition and experience, without possessing the training and theoretical tools needed for this job.

This study investigates lexical, cultural and grammatical problems encountered by Palestinian EFL learners in English-Arabic translation. The findings of this study are intended to draw the attention of those specialists in the field of translation training and syllabus design to the actual translation problems encountered by Palestinian students. The outcomes of this study will contribute to developing the teaching of translation in Palestinian universities in particular, and in Arab universities in general.

Research in translation problems remains ongoing while students continue to commit a wide range of translation errors. A number of studies have discussed translation problems encountered by Arab translators. Some of these have discussed one aspect of the most common translation problems between English and Arabic, e.g. Bahumaid (2006), Al Jabr (2006), Abu Ssaydeh (2004), Zughoul and Abdul-Fattah (2003), Saraireh (2001). Others have discussed different types of English-Arabic translation problems such as Al Kenai (1985), Ghazala (1995), Al Ghussain (2003), Deeb (2005) and Abbadi (2007). Unlike previous studies, the present study will focus on the lexical, cultural and grammatical problems encountered by senior Palestinian EFL learners and view them in terms of their acceptability to qualified experts in Arabic. This linguistic judgment of qualified experts in Arabic will be important in identifying, analyzing and classifying types of lexical, cultural and grammatical problems in the Palestinian EFL students' English-Arabic translations. The study will provide a detailed analysis of the translation strategies employed by the student translators in dealing with various lexical, cultural and grammatical problems. Based on the analysis of three types of texts, namely descriptive, argumentative and instructional, this study will acquaint Arab translators in general and Palestinian translation students in particular with the most challenging translation problems that they might encounter when translating these types of texts.

## **2. Research methodology**

This research, which adopts Kussmaul's (1995: 4) product-oriented approach, consists of three stages: description of errors (looking at the symptoms), finding the reasons for the errors (diagnosis) and pedagogical help (therapy). The subjects chosen for this study were selected from the total population of 500 senior Palestinian EFL learners enrolled in the Department of English, the Islamic University of Gaza in the academic year 2007-2008. Twenty-eight students were selected using the systematic random sampling method. They were 14 male students and 14 female students. All subjects had already studied two translation courses. Students were roughly 20 years old, and their socio-economic backgrounds were similar, since all of them were from the Gaza Strip - a densely populated small area of over a million people - and all of them suffered from the same socio-economic hardships. In addition, the participants' academic level was roughly the same since all of them were admitted to the university according to a fixed graduate point average in their general secondary certificates. The students involved in this study were asked to translate 18 source texts of medium size covering three types of texts: descriptive, argumentative and instructional. The texts, which were selected through close personal monitoring of British newspapers and the BBC, cover topical subjects such as immigration, binge-drinking, the property market, global warming and the Kyoto Protocol, in addition to the Arab Israeli conflict - as the participants were Palestinian refugees living in the Gaza Strip.

In addition to the 18 source texts translated by 28 students from the Department of English, Islamic University of Gaza, the researcher used other instruments of data collection. These instruments are discussed in Sections 3.1 - 3.5.

### **2.1. The researcher's assessment**

The researcher has an MA in Applied Linguistics and ten years of teaching experience at the Department of English, Islamic University of Gaza. He also attended several MA English-Arabic translation modules at the University of Salford, and participated in translation workshops and seminars held at the University of Salford and other British universities. The researcher analysed the translated texts by identifying, describing and explaining the lexical, cultural and grammatical translation problems encountered by senior Palestinian EFL learners. The following is a sample of a student error identification form taken from the translation of source text 3.

**Table 1:** Student error identification form

Student No.	Type of error	Source Text	Target Text	Model Translation
<b>I. Lexical errors</b>				
28	<b>a. Wrong word choice or wrong lexis (unacceptable translations)</b>	It is an all too <u>familiar</u> paradox.  All crime is <u>down</u> 32 per cent over the past decade.  The crime figures show the risk of falling victim to <u>an offence</u> is going down.	على رغم التشابه إلا أنها متناقضة  من هنا فإن كل الجرائم ما دون 32 % تعتبر ضمن العقد الفائت.  الجريمة تظهر الضرر بل الخطر الناجم عن وقوع الضحية في مخالفة قانونية لحد التورط.	إنه لتناقض مألوف جدا.  أن المستوى الكلي للجريمة انخفض بنسبة 32% خلال العقد الماضي.  تظهر أرقام الجريمة انخفاضاً في خطر الوقوع ضحية لهجوم.
	<b>b. Word choice (partially acceptable translations)</b>	The <u>public</u> fear	الخوف العامي	خوف الجمهور
	<b>c. Proper nouns</b>	<u>Wales</u>	وجاراتها في المناطق الجبلية	مقاطعة ويلز
	<b>d. Compounds</b>	But nearly <u>two-thirds</u> of people  <u>Police figures</u>  The <u>authoritative</u> <u>British Crime Survey</u> of <u>offending</u> <u>levels</u>	وبينما هناك حوالي 32% من الناس رجال الشرطة  السلطات البريطانية لدراسة وتفصي الجريمة على كافة مستويات المجتمع	يظل ما يقارب ثلثي الناس أرقام الشرطة  المسح الرسمي للجريمة البريطانية لمستويات الاعتداء
	<b>e. Alternative translation</b>	Do not reflect  The same trend	لم يعكسوا أو يظهروا الحقيقة  نفس النزعة أو الاتجاه	لا تعكس  نفس الميل
	<b>f. Omission</b>	No equivalent for "September"		

	<b>g. Addition</b>	It is not credible to argue that the police figures do not reflect the reality of the situation.	ومن غير المعقول أن نختلف أو نجادل على أن رجال الشرطة لم يعكسوا أو يظهروا الحقيقة في الموقف <u>وذلك على الدوام.</u>	ولا يصح الجدل بأن أرقام الشرطة لا تعكس واقع الحال.
<b>II. Grammatical errors</b>				
	<b>a. Agreement</b>	But nearly two-thirds of people believe crime is increasing.	وبينما هناك حوالي 32% من الناس <u>يعتقد</u> أن الجريمة قد ازدادت.	يظل ما يقارب ثلثي الناس <u>يعتقدون</u> أن معدل الجريمة في ازدياد.
	<b>b. Case ending</b>	But the public fear of crime remains as high as ever.	لكن الخوف العامي من الجريمة يبقى <u>مرتفع</u> نوعا ما	وفي نفس الوقت يظل فيه خوف الجمهور <u>مرتفعا</u> كما كان من قبل.
	<b>c. Word order</b>	The crime figures show the risk of falling victim to an offence is going down.	<u>الجريمة تظهر</u> الضرر بل الخطر الناجم عن وقوع الضحية في مخالفة قانونية لحد التورط	<u>تظهر أرقام الجريمة</u> انخفاضا في خطر الوقوع ضحية لهجوم.

## 2.2. The rater's assessment

The rater involved in the assessment of the translated texts is a native speaker of English and a professor of Arabic at the University of Salford and the supervisor of the author of this study. The rater has long experience in teaching translation and linguistics at British universities, and he has also supervised a considerable number of Arab PhD students. The rater and researcher met on a weekly basis to discuss the translation errors identified by the researcher. The discussion focused on issues related to the identification and categorization of translation errors. In most cases there was unanimous agreement on all issues raised in the discussion.

## 2.3. The target readership assessment

To complement the assessment carried out by the researcher and the rater the translated texts were given to a native speaker of Arabic who has experience in teaching Arabic at university level (cf. Brunette 2000: 173). The native speaker was informed about the main objective of the research and was asked to identify all grammatical and stylistic errors in students' translations. The feedback given by the native speaker of Arabic was so effective that further translation errors were identified.

## 2.4. The professional translator

The professional translator provided a model translation which was used as a model against which students' translations were assessed (cf. Lauscher 2000: 163). The professional translator has long experience in teaching translation at the Islamic University of Gaza. He has also worked as a professional translator in some Arab and international organizations. He is currently doing a PhD in translation and linguistics at the University of Salford.

In assessing the acceptability of the translated texts, a three-point rating scale was used. The rating scale consists of: 'acceptable translation', 'partially acceptable translation' and 'unacceptable translation'. 'Acceptable translations' are judged in terms of meaning and linguistic functions, i.e. translation equivalence conveys the whole intended meaning of the SL message. 'Partially acceptable translations' are also judged in terms of meaning and linguistic functions, i.e. translation equivalence does not exactly convey the whole intended meaning of

the SL message. 'Unacceptable translations', are cases where translation equivalence does not convey the intended meaning of the SL message due to semantic and / or syntactic deviances.

### 2.5. The student questionnaire

The student questionnaire was designed to identify the difficulty of each text. Students were asked firstly to translate the 18 texts, and secondly to rate the difficulty of each text by choosing one of the five options provided under each text (Very difficult, Difficult, Moderate, Easy and Very easy).

### 3. Discussion of results

The results showed that Palestinian students encountered numerous lexical, cultural and grammatical problems when they translated 18 texts covering three types of texts (descriptive, argumentative and instructional) into Arabic. The total number of errors committed across the three types of error was 9253, as shown in the following table:

**Table 2: Summary of students' translation errors found in the study**

Main categories	Sub-categories	No. of errors for female students	No. of errors for male students	Total
<b>A. Lexical errors</b>	Wrong lexis	532	661	1193
	Omission	267	272	539
	Compounds	213	264	477
	Colloquial Arabic	148	255	403
	Transliteration	150	154	304
	Polysemy	111	125	236
	Alternative translation	111	105	216
	Addition	82	132	214
	Collocations	85	88	173
	Acronyms	78	65	143
	Proper nouns	57	59	116
	Metaphors	28	47	75
	Graphic ambiguity	16	24	40
	Idioms	14	19	33
<b>B. Grammatical errors</b>	Punctuation	885	944	1829
	Discourse markers	235	376	611
	Word order	199	350	549
	Case ending	176	190	366
	Referential ambiguity	167	196	363
	Agreement	135	155	290
	Spelling	125	119	244
	Articles	44	89	133

	Tenses	58	71	129
	Prepositions	45	49	94
	Singular and plural	42	44	86
	Passive	30	28	58
	Cataphoric and anaphoric references	0	12	12
C. Cultural errors	Social culture:	165	162	327
	Religious culture			
	Political culture			
	Background knowledge			
	Behaviour			

In the following discussion I will present examples of students' lexical, cultural and grammatical problems. The discussion will also include the results of the gender difference regarding students' translation problems and students' responses to the questionnaire.

### 3.1. Presentation and analysis of lexical problems (4162 errors)

The following discussion of students' lexical problems will focus on the following:

#### 3.1.1. Wrong word choice (wrong lexis)

These errors refer to the semantically wrong word or phrase used by the student translator as an equivalent for the source language word or phrase. These errors produce translations that deviate from the intended meaning of the SL message leading to a distortion of the meaning in the TT. The results revealed that the total number of errors involved in the use of wrong lexis was 1193. To illustrate this point, students gave different translations for the lexical item 'mortgages' in "The number of mortgages approved for homebuyers sank to a record low during October" (Text Seven), as follows:

سجلت معدلات إغراق الرهن العقاري للبايعين المحليين خلال أكتوبر انخفاضاً واضحاً.  
سجل عدد الرهون المبرمة المسقطة لشراء البيوت انخفاضاً في شهر أكتوبر.  
عدد الصكوك يعتمد مع البائعين الغارقين في الديون خلال شهر أكتوبر.

As the above translations show the student translators gave different translations for the lexical item 'mortgage', e.g. الرهن العقاري 'property pledge', الرهون 'pawns, pledges' and الصكوك 'bonds'. These renderings fall short of conveying the intended meaning of the original message. The problem can be related to students' tendency to translate literally, and can also be blamed on English-Arabic bilingual dictionaries which fail to give cultural background knowledge. For example, the two equivalents given by Al Mawrid English-Arabic dictionary for 'mortgage', رهن عقاري and صك الرهن do not convey the intended meaning of the word. The meaning of this word relates to a British economic institution which forms part of British culture. According to the Oxford Advanced Learner's Dictionary, a mortgage is "a legal agreement by which a bank or similar organization lends you money to buy a house, etc; and you pay the money back over a particular number of years".

Some students gave partially acceptable translations, as in:

إن عدد القروض العقارية الموافق عليها لشراء المنازل هبطت لتسجل انخفاضاً خلال شهر أكتوبر.  
أن هناك عدد من موافقات القروض العقارية لشراء البيوت يجب التسجيل في خلال شهر أكتوبر.

عدد القروض العقارية الموافق عليها للمشتريين المحليين غرق إلى هبوط لم يعهد من قبل خلال شهر أكتوبر .

These renderings do not appropriately reflect the intended meaning of the original word since they lack its cultural and situational features. In addition, none of the students provided an explanatory note clarifying the cultural background of the word

### 3.1.2. Omission

Although the translator is obliged sometimes, due to TL-specific conventions or other aspects of untranslatability, to sacrifice some of the items in the SL, omission is not always justified. By 'unjustified' or 'inappropriate' omission, we simply mean omitting certain words, phrases or even sentences that need to be retained in the TL version. This may be attributed to carelessness or failure on the part of the translator to capture the intended meaning and convey it adequately in the TL. The analysis has shown that some students omitted key SL words, phrases or sentences resulting in distorted and non-equivalent translations. There were 558 omissions in the translations. By way of illustration, one student ignored the translation of 'the Home Office' in "According to the Home Office, there were 58 firearms-related homicides in 2006-07 compared with 49 in the previous year-an increase of 18%" (Text One), as follows:

ففي عام 7-2006 هناك حوالي 58 قتل بسبب الأسلحة النارية مقارنة مع 49 حالة في السنة الماضية.

Here, the translator omitted a key phrase in the ST, which is unjustified. This omission not only affects the semantic equivalence but it also reduces the total meaning transferred into the TT. Surprisingly, there is nothing in the context to hinder the translation of this phrase. The whole propositional meaning of the sentence is retained once the omitted part is translated:

و بحسب وزارة الداخلية البريطانية ، تبين أن عدد حوادث القتل باستخدام السلاح الناري لسنة 2006-07 قد بلغ 58 حالة مقارنة ب 49 واقعة سجلت في السنة الماضية، أي بزيادة قدرها 18%.

### 3.1.3. Addition

Unlike omission, addition involves adding one or more words to the source language text in order to make the meaning more explicit for the target language reader. However, addition is not always justified. The results showed that inappropriate additions accounted for 214 errors. For example, one student, translated 'ale' in "Go back into history and there are complaints in Medieval Times about our love of quaffing vast quantities of Ale, and our love of rowdyism and violence" (Text Six), as:

وبالرجوع إلى التاريخ فقد كان هناك في العصور الوسطى يحبون الجعة بشكل واسع وخاصة الجعة الانجليزية الخفيفة وحبهم للفظاظة والعنف.

The target language is unnecessarily repetitious, translating 'ale' as 'الجعة بشكل واسع وخاصة الجعة الخفيفة'. It would perhaps be more appropriate to render 'ale' simply as 'الجعة', as in:

وبالعودة للعصور الوسطى فهناك الشكاوى المتعلقة بحب البريطانيين لشرب كميات كبيرة من الجعة وحبهم للفوضى والعنف.

### 3.1.4. Collocations

Collocation is "a lexical relation of co-occurrence that binds words together with varying degrees of strength" (Bahumaid 2006:133). The most striking examples of collocational difference involve synonyms and by extension near-synonyms. Baker (1992) and Dickins et al. (2002) maintain that synonymous words in English and Arabic often differ in their collocational



ranges, and that each language has its own preferences for certain modes of expression and certain linguistic configurations. This view is also shared by Zughoul (1991: 52), who states that "what collocates in one language does not necessarily collocate in another". The notion of collocation can also be extended in a translation context to cover cases where there is a clear difference in range of meaning between an ST term and its literal TT equivalent. Thus, in addition to its more basic physical sense, 'broken' in English can be used to refer to a society, e.g. 'broken society'. In Arabic, by contrast, the basic literal equivalent of English 'broken' مكسور cannot be used to refer to society; مجتمع مكسور is an unacceptable collocation. The translations of English collocations have shown that students have a preference towards literal translation. As a result they introduced into the target language lexical combinations which are alien to the collocational ranges of their receptor language. The results indicated that the total number of errors found in the use of wrong collocation is 173. To illustrate this point, some students translated 'broken society' in the sentence "Drugs, urban decay, racism, "gangsta" rap and an absence of positive role models have all been blamed for the recent spate of gun deaths, with Tory leader David Cameron going so far as to blame a "broken society" (Text One) , as المجتمع/ المتكسر (المهشم) /مجتمع مكسور as in:

فإن المخدرات، و الفساد الحضري (المدنية)، العنصرية، و قطع الطريق (السطو) و غياب الدور الايجابي للمثل العليا كان سبباً للفيض الأخذ من عمليات القتل بالسلاح، بالنسبة للقائد المحافظ ديفيد كامبرون يذهب للوم ما أسماه "مجتمع مكسور". أدت المخدرات و الانهيار المدني و العنصرية و طرقة رجال العصابات و غياب الدور الايجابي المقندي به إلى الازدياد المفاجئ لحالات الموت باستخدام البنادق حيث أسهب قائد توري: "ديفيد كامبرون" في لوم ما اسماه المجتمع المتكسر / المهشم.

The translations given of 'broken society' are مجتمع مكسور and مجتمع متكسر/ مهشم respectively. These translations are not appropriate to describe the lexical item مجتمع. The translations do not conform to the collocational restrictions of Arabic, i.e. the lexical item مكسور 'broken' never occurs with مجتمع 'society', nor of course, the lexical item مجتمع. The item مكسور occurs with, among other things, material objects such as doors, windows, glass, e.g. باب مكسور 'a broken door', زجاج مكسور 'broken glass'. The lexical item مجتمع co-occurs, inter alia, with منحل فاسد 'corrupt', مفكك 'disintegrated', etc.

A partially acceptable equivalent of the collocation given by one student translator is انحلال المجتمع 'the corruption and disintegration of the society'. However, the same translator wrongly indicated that the Tory leader was behind all the problems inflicted upon British society, as he claimed below:

و لقد أدى هذا كله إلى توجيه أصبع الاتهام و اللوم إلى ديفيد كامبرون و هو قائد حزب المحافظين البريطاني و الذي اعتبر مسئولاً عن انحلال المجتمع وتفككه.

Other students gave acceptable translations of the collocation 'broken society', such as مجتمع مفكك, as in:

السيد المحافظ ديفيد كامبرون ذهب بعيداً للوم "مجتمع مفكك". حيث ذهب القائد في حزب المحافظين البريطاني ديفيد كامبرون إلى أبعد من ذلك فعزى السبب إلى تفكك المجتمع.

The professional translator provided the following acceptable translation:

لدرجة أن السيد ديفيد كامبرون رئيس حزب المحافظين بلغ به الأمر إلى لوم ما أسماه بالمجتمع المفكك.

### 3.2. Presentation and analysis of cultural problems (327 errors)

In addition to lexical problems student translators also encountered cultural problems. The following discussion, which is based on Nida (1964), Ivir (1987), Newmark (1988) and Mailhac (1995), will focus on the following cultural problems:

### 3.2.1. Political culture

Cultures differ in their political terminology (Aziz, 1982: 28). Britain and the Arab World belong to two different political cultures. The former is a democratic country, while the latter is mainly ruled by one-party governments. Thus, political terms used in the two cultures are different, which may constitute a difficulty for the English-Arabic translator. The analysis of translated texts shows that students have difficulty in translating English political terms. This difficulty may be attributed to the differences in the political systems associated with the source and target languages. For example, in Text One some students translated 'Tory leader' in "Drugs, urban decay, racism, "gangsta" rap and an absence of positive role models have all been blamed for the recent spate of gun deaths, with Tory leader David Cameron going so far as to blame a "broken society", as المحافظ القائد 'the conservative the leader', القائد توري 'the leader Tory; and توري كامبيرون 'Tory Cameron', as in:

المخدرات والانهيار المدني والعنصرية وعصابات السرقة وكذلك غياب نماذج الدور الايجابي كانت سبباً في الازدياد المفاجئ لضحايا السلاح ، وذهب المحافظ القائد ديفيد كامبيرون إلى حد بعيد في اللوم بقوله أن ذلك يرجع إلى المجتمع المنهار.

المخدرات، التفسخ الحضاري (انهيار المدنية) والعنصرية وتساعد قوة العصابات مع غياب الدور الايجابي المنوط بالقوانين أو نماذج السلطة، بحيث يفهم سبب هذه الجرائم إلى ارتفاع وزيادة مفاجئة في الجرائم المتعلقة بالأسلحة النارية والقتل حديثاً ( مؤخراً). بينما يتجه القائد توري ديفيد كامبيرون بعيداً عن ذلك من خلال لوم هذا المجتمع المنهار المتصدع.

ويعزى كل من المخدرات والانهيار المدني والعنصرية وعصابات قطاع الطرق وغياب النماذج الايجابية إلى سبل الأسلحة القاتلة مؤخراً. ويصل أيضاً توري كامبيرون إلى حد إلقاء اللوم على المجتمع المهلهل.

Other students omitted the term from their translations, as in:

المخدرات والانحلال المدني والعرقية والعصابات وغياب دور النموذج الايجابي وأيضاً فإن ديفيد كامبيرون... ذهب إلى ابعاد من ذلك ليوم " المجتمع المنشق".

وكانت زيادة معدل انتشار المخدرات ، والتمدن، والانحلال بالإضافة إلى رواج فكرة العنصرية، هذا بالإضافة إلى غياب نماذج الدور الايجابي سبباً للموت بالأسلحة النارية، لذلك أسرع ديفيد كامبيرون... بإلقاء اللوم على المجتمع السائب.

There is a possibility that the translators have dropped these words because they are not familiar with the British political system. The words, however, are relevant to the source text since they are uttered by the leader of the Conservative party, the second largest political party in Britain and the main opposition party at the moment.

### 3.2.2. Background knowledge

The translations show that the majority of students lack cultural background knowledge. Their non-acquaintance with the SL cultural background led some student translators to render some culture-specific features in a literal way, thus missing the cultural significance of the SL text. For example, some students had difficulty in translating the cultural item 'off-licence' in "First, shops and supermarkets must stop selling alcohol below cost prices. This also involves reducing sales to under-18s by bars, off-licences and retailers" (Text Five), as in:

أولاً: على المتاجر والمحلات أن توقف بيع الكحول بأقل من سعر التكلفة. وهذا أيضاً يتضمن تقليل المبيعات لمن هم تحت سن الثامنة عشر من قبل الحانات والغير ممتلكين لرخصة البيع وأصحاب المحلات من بائعي التجزئة.

أولاً: يجب على المحلات والسوبر ماركت أن توقف بيع الكحول بأقل من الثمن المطلوب، وهذا يتطلب تقليل بيعها في الحانات لمن هم اقل من 18 سنة، وكذلك لمن لا تحصل على ترخيص الباعة أيضاً.

أولاً: المحلات التجارية والسوبر ماركت يجب عليهم توقف بيع الكحول من أجل المال. هذا يتطلب أن يقلل من بيع هذه المشروبات للذين أعمارهم تقل عن 18 سنة، والذين لم يتمكنوا من تحصيل الرخصة أو ليس معهم رخصة والبيع بالتجزئة.

It is clear that a literal Arabic rendering of the above instance, e.g. 'Those who do not have a licence to sell alcohol' is not equivalent to the English original. The problem is one of cultural mismatch. Arabic does not have an exact equivalent which covers the whole

semantic field of the word including its cultural aspect. A translation could be accepted only if it is followed by a footnote or an explanatory in-text phrase explaining the contextual meaning of the source cultural item, as in:

أولاً: على المحال التجارية والمتاجر الكبرى التوقف عن بيع المشروبات الكحولية بأقل من سعر التكلفة، ويشمل ذلك الحد من هذه المبيعات لمن هم دون سن الثامنة عشر في الحانات و محلات بيع الخمر (و ليس لتناولها) و محلات البيع بالتجزئة.

(An off-licence is a shop that sells alcoholic drinks in bottles and cans to take away).

### 3.3. Presentation and analysis of grammatical problems (4764 errors)

In addition to lexical and cultural problems, student translators also experienced grammatical problems. The grammatical errors will be discussed under the following sub-categories:

#### 3.3.1. Discourse markers

Discourse markers (such as discourse connectives like 'then', 'so' and 'however') play a major role in binding semantic units and sentences together, serving to achieve cohesion in a text. Arabic and English have different ways of combining sentences. While English makes greater use of subordination, Arabic prefers coordination. In addition, while Arabic favours explicit cohesive markers to link sentences together, English makes common use of both implicit and explicit discourse markers.

The general tendency among students to translate literally resulted in translations that do not conform with the textual forms of the target language. To illustrate, consider the following translation by one of the students:

#### SLT (16): 'Global Warming'

(1) Global warming is one of the biggest problems facing humankind in the next few decades. (2) In the past 150 years, global temperatures have risen approximately 1°C (1.8°F). (3) If temperatures continue to rise, the consequences could be catastrophic. (4) As the earth's temperature rises, polar ice will melt, causing the water level of the oceans to rise. (5) Rising ocean levels, in turn, will cause flooding along the coasts. (6) Global warming will also cause major changes in climate that will affect agriculture.

#### TLT (16):

الاحتباس الحراري العالمي هو أحد المشاكل العظمى التي ستواجه البشرية في العقود القليلة القادمة. ارتفعت درجات الحرارة درجة مئوية واحدة في 150 سنة الماضية. إذا استمرت درجات الحرارة في الارتفاع ستكون العواقب وخيمة /كارثية. إذا ارتفعت درجة حرارة الأرض سينصهر الجليد القطبي وسيسبب ارتفاع في مستوى مياه المحيطات. وفي المقابل، ارتفاع مستوى المحيطات سيؤدي إلى الفيضانات على طول السواحل. وسيسبب الاحتباس الحراري العالمي تغيرات أساسية في المناخ والتي بدورها ستؤثر على الزراعة.

The above translation lacks textual cohesion as some of the sentences are not properly connected. The student's frequent use of full stops has interrupted the flow of the target text. By following the SL structure, the translator failed to convey the implicit textual relations of the SL, creating what, following Baker (1992: 54), can be called calquing, or perhaps better, calquing by omission, showing that the translator is very much engrossed by the SL structure. In other words, when there is an explicit discourse marker in English, the translator often attempts to translate it, but when there is no explicit one, he/she does not insert discourse markers in Arabic. The translator inappropriately used the adversative conjunction وفي المقابل - 'on the other hand' - to join two sentences (sentence 3 and sentence 4) that do not signal such a relation. Failing to identify the logical relationship held between the two sentences, the translator used the wrong connector. The two sentences can be appropriately connected with resultative connectors such as 'as a result', 'consequently', 'so' (لهذا), or 'as a result', e.g.

وعليه فإن ارتفاع مستوى المحيطات سيتسبب بحدوث فيضانات على طول الساحل، وأيضاً سوف يؤدي الاحتباس الحراري إلى تغيرات جوية في المناخ، مما سيؤثر على الزراعة.

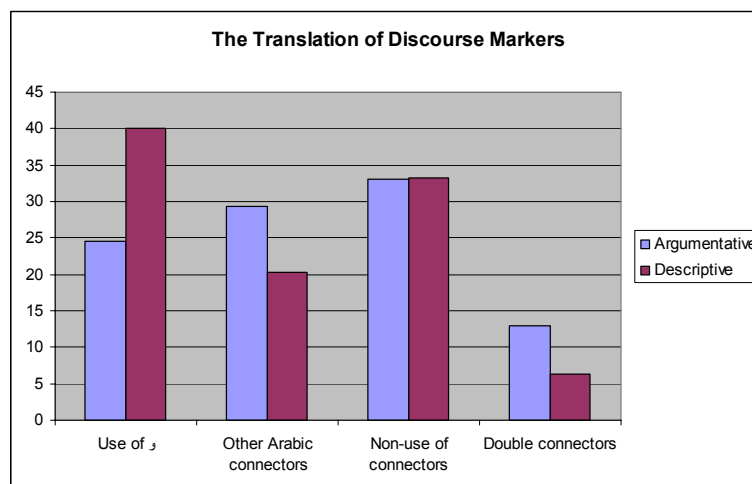
The following tables show students' use of discourse markers in both argumentative and descriptive texts.

**Table 3:** Total number and percentages of use of discourse markers in argumentative texts.

Text	Use of و	Use of other Arabic connectors	Non-use of connectors	Use of double connectors
3	41	42	64	19
6	77	44	77	19
9	53	118	45	29
12	28	28	72	11
15	25	41	43	27
18	27	24	34	27
Total	251	297	335	132
Percentage	24.7%	29.3%	33%	13%

**Table 4:** Total number and percentages of use of discourse markers in descriptive texts.

Text	Use of و	Use of other Arabic connectors	Non-use of connectors	Use of double connectors
1	24	8	23	1
4	18	9	27	2
7	52	12	17	0
10	30	6	19	1
13	10	7	11	0
16	32	42	41	22
Total	166	84	138	26
Percentage	40.1 %	20.3%	33.3%	6.3%



**Figure 1:** Distribution of discourse markers in argumentative and descriptive texts

As indicated in the tables and figure above, students tend to use the additive connector **و** more than other Arabic connectors. In descriptive texts the additive **و** constituted 40.1% of inter-clausal connectors, while in argumentative texts it was recorded at 24.7%. Argumentative texts however differ from descriptive texts in their frequent use of other Arabic connectives (i.e. 29.3% vs. 20.3%). These results coincide with Dickins et al.'s (2002) finding that the general nature of argumentative texts is reflected in the widespread use of connectives expressing logical relations (e.g. **إذن**, **لذلك**, **رغم**, **حيث**, etc.), as well as the more basic connectives, such as **و**, **ف**, and **و**. Argumentative texts also have a higher percentage of double connectors (13% vs. 6.3%). Argumentative and descriptive texts however both show a high percentage of non-use of connectors (33% vs. 33.3%). This tendency to render ST propositions without the use of connectors can be attributed to students' preference for literal translation. It can also be blamed on the failure of some students who lack TL competence and translational proficiency to realize the logical relationships holding implicitly between sentences. Lotfipour-Saidi (1990: 394) has made it clear that "Any careless handling of cohesive elements of the SL text in the translation process would lead to a change in the degree of cohesiveness intended by the original author and such a change would certainly affect the discourse comprehension processes and, in the long run, the nature of the message to be negotiated through the text".

### 3.3.2. Word order

Word order is one of the main features that distinguish one language structure from another. English has subject-verb-object sentence order, while Arabic most commonly has verb-subject-object sentence order. Such sentences, in which the verb precedes the subject, are known as verbal sentences in Arabic. In addition to verb-subject-object word order, verbal sentences may also exhibit verb-object-subject word order. As well as verbal sentences, Arabic also has nominal sentences consisting of a predicand and a predicate. The predicand in nominal sentences is always itself nominal, but the predicate may be a verb (by extension a verb phrase), nominal, adjectival or adverbial. The standard word order in nominal sentences is predicand-predicate, but predicate-predicand word order also occurs. The analysis of translations shows that students have a tendency to use more nominal sentences than verbal sentences in Arabic. This can be attributed to students' preference for word-for-word translation of English elements. The number of errors involving word order (which includes errors related to the overuse of nominal sentences, the use of adverbs or adverbial phrases at the beginning of sentences and the misuse of the emphatic particle **إن**) was 549. To illustrate this point, three students translated "Drugs, urban decay, racism, "gangsta" rap and an absence of positive role models have all been blamed for the recent spate of gun deaths" (Text One), as:

المخدرات و الانحلال المدني و العنصرية و جرم العصابات و غياب دور القذوات الايجابية كل هذه الأمور مسؤولة عن الانفجار المفاجئ الحديث للقتل. المخدرات و انحلال المجتمع ، العنصرية، عصابات الراب و غياب الدور الايجابي للقانون، كلها تلام في ازدياد معدل وفيات السلاح مؤخراً. المخدرات، و فساد المدن، العنصرية، وضجة العصابات، إضافة إلى غياب المبادئ الأخلاقية الايجابية كل هذه العوامل أدت إلى ازدياد نسبة القتل المسلح.

In the following translations students kept the source language word order, producing a shift in sentence focus. By opening with the adverbial phrase **خلال السنة الماضية**, the student translator has moved the emphasis to the adverbial phrase 'during the last year'. The occurrence of the adverbial phrase **خلال السنة الماضية** affects the style of the text, producing a style which does not fit the structure of the TT.

#### SLT

"During the past year the deliberate use of guns to take life has risen in England and Wales" (Text One), as follows:

#### TLT

خلال السنة الماضية ارتفع الاستخدام المتعمد للسلاح للدفاع عن النفس في انجلترا وويلز. خلال السنة الماضية فإن الاستخدام المتعمد للسلاح للقتل قد ازداد في انجلترا وستراليا.

### 3.3.3. Passive

In translating English passive sentences into Arabic, student translators used several techniques. They include:

1. Using other Arabic forms (e.g. Arabic nominals and active participles) (39.1%).
2. Arabic passive without a 'by'-type phrase (25.4%).
3. Translating by omission (14.7%).
4. Using passive-like forms (e.g. passive participles) (7.3%).
5. Using Arabic active (6.5%).
6. Dummy verbs such as تم plus a verbal noun (quasi-passive) (4.9%).
7. Arabic passive plus an agentive phrase (2.1%).

The use of several techniques in translating English passive into Arabic has shown that Khafaji (1996: 23) is right when he says: "The relatively free word order in Arabic, in addition to its rich verb morphological system, provides it with alternative means of expressing passivity other than by using the morphologically passive verb". However, the translation of the passive has revealed erroneous translations due to students' lack of awareness of the passive system in English and Arabic and their unfamiliarity with the various options available to them for translating the English passive into Arabic. For example, some students translated English agentive passive sentences into Arabic passive sentences in which the agent is overtly expressed. One instance of this occurred where some students translated the English agentive by-phrase in "The number of mortgages approved for homebuyers sank to a record low during October, according to figures released by the British Bankers' Association" in Text Seven, as follows:

انخفض عدد الرهانات العقارية المتاحة للمشتريين المحليين ليسجل انخفاضاً خلال شهر أكتوبر، وفقاً للمؤشرات التي أجريت عن طريق الرابطة البريطانية لأصحاب البنوك.

عدد القروض العقارية التي تمنح لمشتري البيوت انخفضت إلى رقم قياسي خلال أكتوبر، حسب الأرقام التي نشرت بواسطة جمعية البنوك البريطانية.

Although Modern Standard Arabic has come to exhibit examples of agentive prepositional phrases such as 'from the side of', 'at the hands of', 'by means of', and 'on the part of', due to the influence of translation (Badawi 2004: 385), Arab translators should be familiar with the similarities and differences in the voice systems of English and Arabic in order to avoid using constructions that may weaken their Arabic style.

### 3.4 Gender differences regarding students' translation errors

The results showed that male students committed more errors than female students. The total number of errors committed by male students is 5055, compared with 4198 errors committed by female students. Grammatical errors show the biggest gender difference: male students committed 2623 errors, while female students committed 2141 errors. However, male students made fewer cultural errors than females. While male students committed 162 cultural errors, female students committed 165.

### 3.5. The results of student translators' responses to the questionnaire

In this study the student translators were asked to underline the difficulty they encountered in translating the texts by choosing one of the five options given (i.e. Very difficult, Difficult, Moderate, Easy and Very easy). The results revealed that argumentative texts were regarded by students as the most difficult at 44%, followed by instructional texts at 26.2% and finally descriptive texts at 21.4%.

These results accord with the errors detected in this study. As can be seen in Table 9 below, argumentative texts were the most difficult at 45%, followed by instructional at 28.8% and finally descriptive texts at 26.2%. The results of the questionnaire also showed that Text 18 was regarded by students as the most difficult, while the analysis of students' translations revealed that Text 9 was the most difficult (924 errors). However, both the questionnaire and the analysis showed that Text 10 was the easiest text (248 errors). This can be attributed to students' familiarity with such types of texts as they, as Palestinians, have been living under the occupation for more than five decades.

**Table 5:** Students' lexical errors per text

Text	WL	OM	CO	CA	TR	PO	AT	AD	CL	AC	PN	ME	GA	ID	Total
1	58	20	31	19	0	0	16	14	34	0	6	0	0	0	198
2	93	30	66	25	1	1	27	16	19	0	0	0	1	0	279
3	67	22	36	18	1	5	18	15	6	0	0	0	0	0	188
4	46	15	2	40	0	0	10	9	11	0	0	0	0	0	133
5	100	26	28	32	9	2	13	5	1	0	0	1	1	0	218
6	46	10	7	36	0	1	13	13	8	7	6	22	2	0	171
7	66	45	74	47	42	52	11	7	14	18	0	17	1	0	394
8	47	16	2	22	65	7	9	7	2	0	0	0	1	11	189
9	70	29	33	48	26	21	30	14	3	0	25	16	1	22	338
10	53	11	0	4	0	3	7	8	0	0	24	0	0	0	110
11	82	41	9	6	0	18	8	12	9	30	0	0	10	0	225
12	92	54	9	12	0	7	6	29	12	21	10	8	9	0	269
13	49	13	2	10	0	18	4	5	6	15	0	0	0	0	122
14	64	29	40	21	1	4	10	16	27	0	45	0	2	0	259
15	70	43	11	18	20	3	3	6	3	0	0	11	4	0	192
16	27	29	0	10	43	0	9	15	4	0	0	0	1	0	138
17	72	35	55	24	31	34	7	11	5	22	0	0	3	0	299
18	91	71	72	11	66	60	14	12	9	30	0	0	4	0	440

**Key to errors:** **WL** Wrong Lexis, **OM** Omission, **CO** Compound, **CA** Colloquial Arabic, **TR** Transliteration, **PO** Polysemy, **AT** Alternative Translation, **AD** Addition, **CL** Collocation, **AC** Acronyms, **PN** Proper Nouns, **ME** Metaphor, **GA** Graphic Ambiguity, **ID** Idioms.

**Table 6:** Students' grammatical errors per text

Text	PU	DM	WO	CE	RA	AG	SL	AR	TE	PR	SP	PS	CA	Total
1	14 9	20	33	21	0	7	10	0	0	1	7	0	0	248
2	12 9	0	6	13	0	0	11	6	0	1	1	0	0	167
3	85	78	36	12	0	9	23	4	2	6	4	1	0	260
4	81	47	37	24	0	23	10	8	13	9	0	5	0	257
5	11 5	0	16	24	0	10	14	2	0	6	1	0	0	188
6	15 8	94	26	30	12 0	45	22	4	3	4	4	2	0	512
7	84	28	17	35	0	13	33	23	11	7	3	7	0	261
8	95	0	0	8	0	6	10	2	3	7	10	2	0	143
9	16 2	69	45	22	78	39	14	5	53	7	8	1	0	503
10	56	27	5	14	0	12	6	4	3	2	9	0	0	138
11	11 8	0	8	23	21	5	8	2	0	4	0	1	0	190
12	18 1	47	42	19	54	25	15	29	8	7	2	0	12	441
13	49	10	24	39	0	17	9	0	7	6	1	0	0	162
14	54	0	53	28	6	28	22	19	1	4	5	0	0	220
15	77	67	32	27	83	36	11	9	11	4	2	3	0	362
16	41	63	43	15	0	6	4	1	1	4	2	0	0	180
17	72	0	58	2	0	0	8	1	1	5	3	28	0	178
18	12 3	61	68	10	1	9	14	14	12	10	24	8	0	354

**Key to errors:** **PU** Punctuation, **DM** Discourse Markers, **WO** Word Order, **CE** Case Endings, **RA** Referential Ambiguity, **AG** Agreement, **SL** Spelling, **AR** Articles, **TE** Tense, **PR** Prepositions, **SP** Singular and Plural, **PS** Passive, and **CA** Cataphoric and Anaphoric.

**Table 7:** Students' cultural errors per text

Text	SC	PC	BK	Total
1	5	23	28	56
2	0	0	0	0
3	0	0	0	0



4	0	0	0	0
5	28	0	81	109
6	51	0	0	51
7	28	0	0	28
8	0	0	0	0
9	0	0	83	83
10	0	0	0	0
11	0	0	0	0
12	0	0	0	0
13	0	0	0	0
14	0	0	0	0
15	0	0	0	0
16	0	0	0	0
17	0	0	0	0
18	0	0	0	0

**Key to errors: SC** Social Culture, **PC** Political Culture, **BK** Background Knowledge.

**Table 8:** Total number of errors per text

Text type	Number of texts	Number of errors	Percentage
Descriptive	1,4,7,10,13,16	2425	26.2%
Instructional	2,5,8,11,14,17	2664	28.8%
Argumentative	3,6,9,12,15,18	4164	45%
	Total	9253	100%

### 3.6. Conclusion

The results showed that Palestinian students encountered numerous lexical, cultural and grammatical problems when they translated 18 texts covering three types of texts (descriptive, argumentative and instructional) into Arabic. These errors are attributed to the following factors:

1. A general tendency among student translators to translate literally.
2. Students' dependence on existing bilingual English-Arabic dictionaries such as Al Mawrid English-Arabic Dictionary.
3. Lack of linguistic competence and translational skills such as proof reading skills.

4. Students' unfamiliarity with target language culture and lack of knowledge of some aspects of the target culture.
5. Students' unfamiliarity with text types and the two notions of monitoring and managing.
6. Lack of English-Arabic specialized dictionaries such as dictionaries of collocations and idioms.
7. Students' unfamiliarity with the similarities and differences of lexical, cultural and grammatical features in English and Arabic.

## References

- Bassnett, S. (1980). *Translation Studies*. London: Routledge.
- Catford, J. (1965). *A Linguistic Theory of Translation*. London: Oxford University Press.
- Dickins, J., Herve, S. and Higgins, I. (2002). *Thinking Arabic Translation: A Course in Translation Method: Arabic to English*. London and New York: Routledge.
- Dickins, J. and Watson, J. (1999). *Standard Arabic: An Advanced Course*. Cambridge: Cambridge University Press.
- Dollerup, C. and Loddegaard, A. (1992). *Teaching Translation and Interpreting. Training, Talent and Experience*. Amsterdam: John Benjamins.
- Fawcett, P. (1997). *Translation and Language: Linguistic Theories Explained*. Manchester: St Jerome.
- Hatim, B. and Mason, I. (1990). *Discourse and the Translator*. London and New York: Longman.

# Exploration Study of the Fishing Industry of Libya

**Nagat Elmsallati, Prof. David Eaton**, *School of the Built and Human Environment, University of Salford*

## Abstract

This paper presents the main factors in establishing an industrial scale fishing industry in Libya with respect to the types and location of principal landing sites along the coast. The types and distribution of fishing fleets working over the coastline, such as artisanal fishing fleets represented by canoes, small gill-netters (flouka), large gill-netters (Motor), light fishing (lampara), local and overseas trawlers and fish-factory ships are examined with reference to the current situation for fish production, fish exporting and marketing facilities. In addition to this, solutions to the problems facing fish production and marketing with respect to handling by fisherman and consumers are presented. The conclusion of this paper includes some scientific and applied recommendations to develop such an important economic activity in Libya.

## 1. Introduction

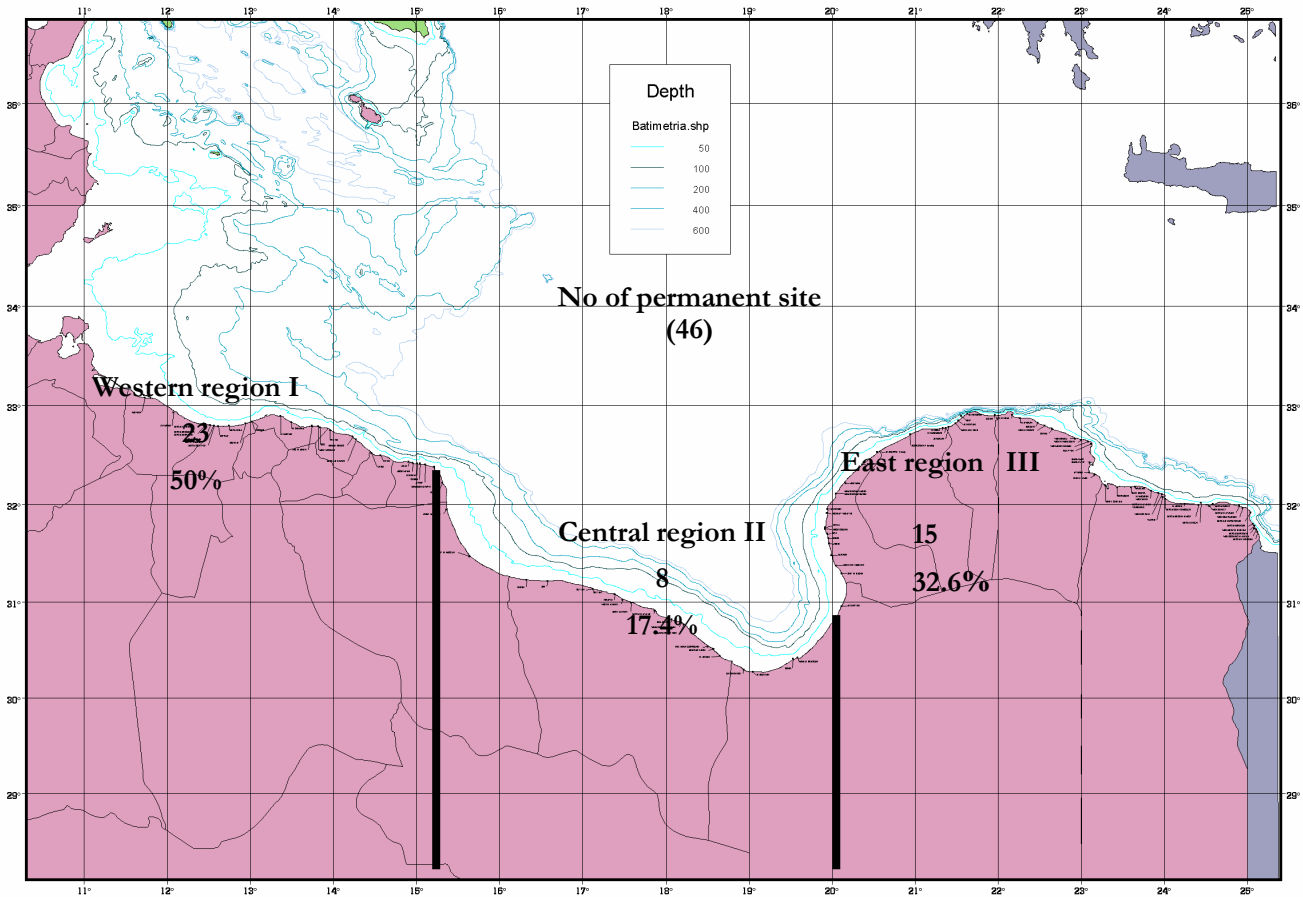
Libya is an Arab country of North Africa. It is bounded by the Mediterranean Sea on the north, where the capital is Tripoli. The total length of the Libyan coast is approximately 2000 km, it is one of the longest coastlines in the southern Mediterranean and one of the most important fishing areas (Malvarosa, 2002). The Libyan coast is divided into three main areas, which are East, Central, and West, according to the depths along the Continental Shelf. The borders of these three divisions are as follows: the Eastern region - extending from the Libyan-Egyptian border to Azwaitinh; the Central region - stretching from Azwaitinh to Marzukh; the Western region - stretching from Marzukh to the Tunisian border to the west.

## 2. Landing Sites

A landing site is a harbour or shelter or other site protected from sea waves and safe for the anchoring of the fishing fleets, as well as the place where the production of the fishing fleets is landed. The landing sites vary according to the nature, forms, and their usage during the year as permanent or seasonal fishing for a particular type of fish (Malvarosa, 2002).

There are 135 landing sites along the Libyan coast, 46 of these sites are permanent (all year operations) and 89 are seasonal. The distribution of these sites on the Libyan coast is 69 (51.1%) in the Eastern region, 32 (23.7%) in the Central region and 34 (25.2%) in the Western Region (Marine Biology Research, 2008).

Despite the Eastern region recording the largest number of landing sites (69), the Western region recorded the largest number of permanent sites - 23 (50%), whereas the Eastern region and Central Region recorded 15 (32.6%) and 8 (17.4%) respectively of the total number of permanent landing sites as shown in Figure 1.



**Figure 1:** Distribution of permanent landing sites along the Libyan coast (Alzkose, 2005).

### 3. Fishing fleets

Fishing activities in Libya can be classified into two main categories, namely: artisan fishing fleets and industrial fishing fleets (Table 2).

Artisan fishing fleets are represented by canoes, small gill-netters (flouka), large gill-netters (Motorised), and Lampara fishing (a technique using lamps to attract the fish),

Canoes are boats with flat bottoms of 7 – 8m length. They are generally powered by out-board engines and operate with gillnets, trammel nets, longlines, and handlines. Some canoes in Farwa lagoon are used to fish octopus using pots.

Small gillnetters (Flouka) are boats of 2 – 7m in size, tapered at both ends, or with a flat transom. All are equipped with out-board engines. These small boats fish with gillnet, trammel nets, and longlines.

Large gillnetters (Motorised) are boats generally greater than 5 – 18m with roofed wheel house, net hauler, and fish hold. These craft operate with gillnets, trammel nets and longlines.

The Lampara fishing system consists of two boats, the main boat is usually 12 – 13 m (maximum 18m) in length and has a fish-deck, inboard engine, a small enclosed cockpit, and a purse seine winch. The second boat is 7 – 8 m in length, has no deck or engine, and carries kerosene lights. These units fish with purse seine nets and use the light to attract the fish at night, catching sardine, and mackerel. The lampara fishing season lasts from April to September/October. During the rest of the year some of these boats are converting to net or line fishing. Lamparas are only present in the western part of Libya (Table 3) (Lamboeu *et.al.*, 2006: 5- 9).

The industrial fishing fleet is represented by coastal trawlers, overseas trawlers and ships, and the Tuna fishing fleet. It is composed (excluding the Tuna fishing fleet - for which there are no available statistics) of 123 units (Table 4), which are steel and wood stern trawlers, mostly owned privately, either by individuals or partnerships.

This type of fishing fleet is completely different from the artisan fleets, using fishing trawler-nets. With a high level of technical specification, given its size and capabilities of the equipment, devices, storage capacity, cooling and processing of fish. It is divided into two types:

1) Coastal trawlers are fishing units of specific technical specifications, large and medium sized and the fishing operations are in coastal water by special nets and the number of fishermen per boat consists of 8-12;

2) Overseas trawlers and ships with high technical specifications of large size, storage capacity, cooling, processing of fish and high capacity fuel tanks that enable them to carry out fishing on long journeys. This kind of industrial fishing fleet has a crew size from 12-25 people (FAO, 2007).

The third type of industrial fleet is the dedicated Tuna fishing fleet. Blue fin tuna (*Thunnus Thynnus L.*) is a large pelagic fish and highly migratory, it migrates every year from the Atlantic through the Straits of Gibraltar to the Mediterranean. The route of migration usually takes place along the African coast (Morocco, Algeria, Tunisia, Libya), most of the fish are large sized, mature fish that are due to spawn in the Mediterranean during early summer. Taking this opportunity the fish are caught by several methods (FAO, 1999)

Tuna fishing is carried out mainly by the industrial fishing fleet of longliners and purse seiners, and tonnaras (a set of nets extending 3-5 km out from the coast). The longline used to catch bluefin tuna consists of six fishing vessels. The length of the line using to catch blue fin tuna is about 100-120 km. and the number of hooks between 2500 - 3000. The fishing area is in the west part of the Gulf of Sirte. At some times the production of bluefin tuna was exported directly without landing in any of the ports. However the production from the tuna traps was canned locally immediately after catching (FAO, 1999).

Table 1 shows the production of bluefin tuna (BFT) throughout 2000 – 2007. The total catch of BFT from the Libyan waters is in metric tonnes.

**Table 1:** Production of BFT (2000 – 2007)  
Source: General Marine Wealth Authority (2008)

Year	Production
2000	19617
2001	21567
2002	19831
2003	22375
2004	38514
2005	27107
2006	39215
2007	33345

#### 4. Catch utilization

Most of the catch is taken by the artisan fishing fleets, and all Libyan catches are presently sold fresh from large urban market areas (except for a part of the small pelagic stock – mackerel and sardine -which goes to canning for the domestic market or as a fishmeal, during high peak production).

Seven fish canning plants belonging to State companies exist which can process tuna and small pelagic stock, with a daily (raw material) capacity, 85 tonnes of tuna, 51 tonnes of small pelagic and 130 tonnes of fishmeal. None of these plants seem to operate in a satisfactory condition due to problems in supply of raw material and the poor state of equipment in some of them. Although the government decided to engage in a broad privatization policy which aimed to progressively withdraw the State from the entire productive sector, privatization is still underway for these canning plants since 2003 (NEPAD, 2008; FAO, 2007).

In general, exports of fish products are still very low. About 2,000 tonnes of BFT are exported annually to international markets (mostly to Japan) and small quantities around 3,000 tonnes of high value fish, are exported to neighbouring countries mostly to Tunisia (NEPAD, 2008; FAO, 2007). In general, the export of fish products is still very limited, due to gaps in legislation covering requirements adequately, a laboratory network for sampling of fish products, and an upgrade of export facilities up to international standards (FAO, 2006).

**Table 2:** Number of fishing vessels  
Source: Marine Biology Research Centre, 2008

Year	Artisan	Industrial
1994	2855	105
2000	1866	105
2002	1895	65

**Table 3:** Distribution of artisan fleet along the coast (2002)  
Source: Marine Biology Research Centre, 2008.

Region	Canoes	flouka	Motor	Lampara	Total	Percent
Western	64	788	343	129	1324	69%
Central		102	55		157	8%
Eastern	13	250	167		430	23%
<b>Total</b>	77	1140	565	129	1911	
<b>Percent</b>	4%	60%	29.5%	6.5%		

**Table 4:** Distribution of Libyan industrial fleet along the coast (2002)  
Source: Marine Biology Research Centre, 2008.

Region	Landing site	Trawlers	Total	Percent
Western	Zwara	16	48	77.4%
	Bab albahar	21		
	Al shaab	5		
	Al nadi al bahari	1		
	Mena al khoms	5		
Central	Mena ras lanof	1	1	1.6%
Eastern	Mena bengazi al gaded	11	13	21%
	Mena bengazi al gadem	1		
	Mena tuburk	1		
<b>Total</b>		62	62	

## 5. Methodology

This was the first occasion to be directly in contact with the fisheries environment of Libya. Thus a qualitative heuristic exploratory approach was adopted. The researcher chooses two investigative methods in this study including: direct observations and semi-structured interviews. The researcher tried to build a rapport with the respondents in order to understand the respondent's point of view rather than make generalisations about behaviour.

The semi structured interview technique was used to collect qualitative data by setting up a situation (the interview) that allows a respondent the time and scope to talk about their opinions on a particular subject. This type of interview (semi-structured) is flexible, allowing new questions to be brought up during the interview as a result of what the interviewee says (Oppenheim, 2000). The interviewer in a semi-structured interview generally has a framework of subjects to be explored. Thus; the researcher conducted the interviews with professionals involved in the marine wealth sector, researchers from the Marine Biology Research Centre and with some fishermen.

Observational tools have also been used as a supportive or supplementary technique to collect data. An observation can take a variety of forms and can be used for a variety of purposes (Collis and Hussey, 2003). In this study the researcher used the observational tool in the form of photography. A set of photographic interventions in several locations inside the market and in some permanent and seasonal landing sites was conducted, these photographs have been used to validate the discussions which have been made and to triangulate the information obtained from the interviews,

## **6. Findings and conclusions**

Through previous research literature in the marine wealth sector, particularly with regard to the problems and difficulties of fishermen, vendors and consumers, and through discussions with officers and researchers in this field, in addition to some field visits and interviews with employees, the Libyan problems have been summarized as:

- Fishing boats are old;
- Non use of modern technology in fishing;
- Most ports are unfit to sail from;
- Clear evidence of pollution in the sea water;
- No services are available at the harbours for the fishermen (e.g., Toilet and bathrooms);
- Most labour is foreign and transient;
- Libyan labour has no desire for this job;
- The condition of fish markets is very poor;
- Fish display was very unhygienic;
- There was a lack of channels for marketing fish, especially Lampara fish;
- Low selling prices compared with the price of ice (which is considered as the most important in controlling the spoilage of fish);
- Problems in the conduct of production and consequently the need for specialized markets for selling fish since seafood needs special treatment, in handling, transportation and distribution from production centres to shopping centres, because of its quick spoiling;
- Production losses causes big losses to the fishermen and dealers;
- Most markets are available only near production centres so this creates difficulties to reach the consumers in many areas.

In conclusion, the proposals for solving such problems can be by constructing facilities for providing services to the fishing industry. This includes the erection of marine workshops specializing in boat and engine maintenance and establishing practical training centres for the fishing industry and encouraging national labour to join them. Another solution would be opening efficient modern shops and markets for selling fish all over Libya and establishing more factories for processing Lampara fish; particularly for sardines caught in large quantities during the summer and frozen and sold during winter and spring. Organizing training and knowledge courses for fishermen and all the workers at the harbours is required for the application of the Hazard Analysis and Critical Control Point system (HACCP) into the fisheries industry to improve the quality of the fish sector in Libya. These measures will contribute to the development of a viable and sustainable industrial fishing industry for Libya.

## References

- Alzkose, S. (2005), *Study the status of fisheries and fishing in Libya at the period of 2002-2004*. Marine Biology Research Centre (MBRC), Tajura- Tripoli Libya.
- Collis J and Hussey R (2003), *Business research: a practical guide for undergraduate and postgraduate students*. Palgrave Macmillan, Basingstoke.
- Food and Agriculture Organization of the United Nations, Fisheries and Aquaculture Department (1999) *Research on Fishing Biology of Blue Fin Tuna (*Thunnus thynnus*) and sword fish (*Xiphias gladius*) in the Mediterranean sea*, Available at: [http://www.faocopemed.org/old\\_copemed/reports/tunidos/tunidos00/manuales/estandarising.pdf](http://www.faocopemed.org/old_copemed/reports/tunidos/tunidos00/manuales/estandarising.pdf).
- Food and Agriculture Organization of the United Nations, Fisheries and Aquaculture Department (2007), *Fishery & Aquaculture Country Profile, Libyan Arab Jamahiriya.*, Available at [http://www.fao.org/fishery/countrysector/FI-CP\\_LY/en](http://www.fao.org/fishery/countrysector/FI-CP_LY/en).
- Hussey J. and Hussey R. (1997), *Business research: a practical guide for undergraduate and postgraduate students*, Macmillan Press, London
- Lamboeul M., Ben Abdallah., A., Amer, A., Turki, A., El-Gushti, A., Fituri, A., El-Srif, E., Shugman, E, Ghebli, H., and Ghmati, H., (2006), *Artisanal fisheries in Libya, Census of fishing vessels and inventory of artisanal fishery metiers.*, *fao- copemed-mbrct*, Food and Agriculture Organization of the United Nations (FAO)
- Malvarosa, L. (2002), "The fish trade of North African Mediterranean countries: intra-regional trade and import export with the European Union", *FAO Fisheries Circular*. No. 978, Rome, FAO.
- MBRCT (2008) "Annual report of Marine Biology Research Centre" Tajura Tripoli Libya
- NEPAD (2008), *New Partnership for Africa's Development, Government Of The Libyan Arab Jamahiriya*, Fisheries Development Project, Volume V: Fisheries Development Project.
- Oppenheim A. N. (2003), *Questionnaire design: interviewing and attitude measurement*. Printer, London.
- Reynolds, IJE. Abukhader, A. & Ben Abdallah, A.(1995), *A Development Planning Overview, The Marine Wealth Sector Of Libya*, Libfish Field Documents, No. 14, Technical Assistance In Fisheries Development, Food and agriculture organization of the united nations, Rome.



# The Impact of Contingent Factors on the Use of Performance Measurement Systems in the Banking Industry: The Case of Libya

**Gumma Fakhri, Roger Pegum, Karim Menacere**, *Faculty of Business and Law, Liverpool John Moores University*

## Abstract

The aim of this paper is to investigate the application of financial and non-financial performance measures in the Libyan banking sector and to assess to what extent the Libyan banks adopt modern performance measurement techniques such as balanced scorecard (BSC). In addition, the paper will examine the impact of contingent factors on the use of financial and non-financial performance measures from a contingency perspective. Thus, the study identifies the reality of performance measures taken from the performance measurement literature, and investigates the effect of four selected contingent factors: namely, organisational structure, level of competition, size of the bank, and business strategy of banks on the use of performance measures. Based on a scale survey in a sample of 68 respondents from diverse banks in Libya, the study develops hypotheses concerning the study objectives, and employs descriptive analysis to assess the underlying impact of using financial and non-financial performance measures and evaluate the adoption of BSC. Furthermore, bivariate correlation statistics are utilised to assess the impact of selected factors on the use of performance measures.

The study finds that most Libyan banks are still relying heavily on financial measures, even if they tend to place strong emphasis on customers and the quality dimensions of non-financial measures. Moreover, few Libyan banks are adopting BSC, and they tend to implement customer-related measures and learning and employee growth measures more frequently than banks that do not adopt them. The study also finds that all four contingent factors have varied positive impact on the use of performance measures in Libyan banks.

**Keywords:** performance measures, performance measurement systems, Libyan banking sector, contingency theory

## Introduction

Our review of the management accounting literature underlined how important management accounting functions are for providing appropriate information about internal activities and assuring the financial success of the organisation (Drury, 2004). Therefore, firms are focused on the use of management accounting information to help managers make basic decisions in order to achieve organisational objectives. The performance measurement system is one of the most important functions of management accounting, as it is employed to evaluate, control, and improve processes through comparing the performance of different organisational levels (Drury, 2004). So it is important for managers to both track and measure performance for their subunits, and for employees at lower levels to understand the financial impact of their operating decisions (Anthony and Govindarajan, 2001). Consequently, it is essential to understand that performance measurements are used at each organisational level. This paper focuses namely on financial and non-financial performance measures.

## A brief literature review and development of the hypothesis

The literature we reviewed considers the use and limitations of financial measures on performance evaluation. Recent changes in the business environment—such as technological advances and increased competition—pose challenges for organisations. This has led management accounting specialists to criticise the exclusive reliance on financial measures for performance assessment (e.g., Atkinson et al., 1997; Ittner and Larcker, 1998; Neely, 1999). As a result, the literature recommended that organisations ought to use non-financial as well as financial measures to provide managers with adequate information about their overall

organisational performance (Kaplan and Norton, 2001; Banker et al., 2000). In addition, contingency theory literature suggests that the organisation's business environment should influence the shape of management accounting practices, including performance measurements (Otley, 1980; Ittner and Larcker, 2003). In line with this proposal, Otley (1999) concluded that performance measurement systems used in one company may not be suitable for another company facing different circumstances. Therefore, the following sections will briefly consider relevant theoretical and empirical literature regarding the use of financial and non-financial measures for performance assessment, together with the contingent factors that might affect the use of performance measures:

### **Financial measures**

Financial performance measures are employed to provide financial information to managers and other users, and to evaluate efficiency and effectiveness. The more popular financial measurements include return on investment, return on assets, return on capital employed, and earnings per share (Ittner et al. 2003). Although financial performance measures are important in performance assessment, they have some limitations. For example, Kaplan and Norton (1996) and Neely (1999) agree that financial measures are too financially oriented, internal looking, focussed on inputs rather than outputs, and short-term oriented. These limitations indicate that financial measures should be expanded to incorporate the valuation of the company's intangible and intellectual assets. The latter include high-quality products, motivated and skilled employees, responsive and predictable processes, and satisfied and loyal customers, which are all critical for success in today's competitive environment (Kaplan and Atkinson, 1998). These types of measures can be categorised as non-financial performance measures. Furthermore, Kaplan and Norton (1996) argue that using financial measurements exclusively can damage an organisation's capacities. They recommend that a combination of financial and non-financial measures is better suited for evaluating performance.

### **Non-financial measures**

Several recent studies have provided empirical evidence of the positive impact of non-financial performance measures on organisations' long-term financial performance (Anderson and Lanen, 1999; Banker et al., 2000). Non-financial performance measures provide managers with timely information concentrated on the causes and drivers of success and can be used to design integrated evaluation systems (Fitzgerald et al., 1991; Kaplan and Norton, 1996; Banker et al., 2000). Fisher (1995) states that there are three main reasons for the emergence of non-financial performance measures: the limitations of traditional financial performance measures, competitive pressures, and the growth of other initiatives. In addition, Neely (1999) highlights several reasons for this performance measures revolution, including increased competition, changing organisational roles, changing external demands, and the power of information technology. This in turn has led to the recognition that financial performance measures do not present a clear picture of organisational performance (Bourne and Neely, 2002). Most studies of non-financial performance measures have focussed on manufacturing; very few concern services firms (Kald and Nilsson, 2000). Several studies (Fitzgerald et al., 1991; Kaplan and Norton, 2001; Hussain, et al., 2002; Lorenzo, 2008) have emphasised the need to use multidimensional performance measures in the service sector—for example, in banking. In their discussion of UK bank-lending decisions, Berry et al. (1993) argue that although manufacturing companies tend to emphasise the importance of non-financial performance measures, bankers are more concerned with financial performance measures. Ostinelli and Toscano (1994) examine the use of two non-financial measures—customer satisfaction and improvement in quality management—as operational tools of control in three Italian banks. They found that the management control system was able to integrate both financial and non-financial measures to evaluate performance. Hussain et al. (2002a) researched the role of management accounting practices in non-financial performance measures in financial institutions (including banks) in three countries: Finland, Sweden, and Japan. Their study found that economic, normative, coercive, and other contextual factors have affected the role and the use of non-financial performance measures in the financial sector in three different countries. Al-Enizi et al. (2006) examined the use of non-financial performance measures in four service companies (including a bank) in Gulf Cooperation Council countries. They suggested that non-financial performance measures have a positive impact on long-term profitability. Hussain and Hoque (2002) examined factors affecting the

design and use of non-financial performance measurement systems in Japanese banks. They found that several institutional features influenced the banks' decision to implement a particular performance measurement system, including the central bank's regulatory control, bank size, and competition. Hussain and Hoque (2002) assessed the role of management accounting in non-financial performance among Japanese financial institutions. They reported that management accounting has played a key role in measuring performance in different banks in Japan, but noted that its role in non-financial performance measures has been less significant than its role in financial performance measures. The authors concluded that contextual factors affected the use of non-financial performance measures and that non-financial performance measures are needed.

The above discussion suggests that a few empirical studies have directly examined the use of financial and non-financial measures for performance evaluation in the banking industry in some developing countries, but not in Libya. In addition, related studies provide two main arguments regarding the use of financial and non-financial measurements. The first points out that the use of financial measures is more common and standardised than the use of non-financial measures across the organisation's subunits, since financial outcomes are the primary performance objectives. The second argument concludes that non-financial measures are of great utility than financial measures in performance assessment systems. The former, it is claimed, are better suited to drive future financial performance, and they reflect the value of long-term assets. Over the last decade, the balanced use of financial and non-financial measures for performance assessment has been strongly recommended by scholars and professionals (e.g., Kaplan and Norton, 1996).

On the conviction that financial measures are critical for performance evaluation in the Libyan banking sector context, this paper sets forth its first hypothesis:

***H.1:** Libyan banks tend to utilise financial measures more frequently than non-financial measures.*

### **The balanced scorecard technique**

The balanced scorecard (BSC) is a fashionable performance measurement system that was introduced by Kaplan and Norton (1992). It incorporates both financial and non-financial measures. In the version of BSC designed by Kaplan and Norton (1996), four aspects of performance measurement are emphasised. The first—the financial perspective—addresses the ways in which companies can maximise value creation for shareholders. Second, the customer perspective helps the company to meet customers' requirements and maintain their loyalty. Third, the internal business process perspective address practises that can help the organisation gain competitive advantages. Finally, the learning and innovation perspective helps the organisation to determine how to maintain its ability to change and improve. Therefore, the key advantage of applying BSC is that allows an organisation to utilise an aggregation of financial and non-financial performance measures, which enable it to track financial outcomes and monitor progress in building the capabilities needed for growth (Kaplan and Norton, 1992; Ittner and Larcker, 1998; Hoque and James, 2000; Guenther and Gruening, 2002). On the other hand, Neely et al. (2001) indicate that the balanced scorecard has several limitations. For instance, it does not consider the competitor perspective, and it is not applicable to all organisation levels.

The foregoing discussion suggests that the adoption of BSC should improve the use of financial and non-financial performance measures. The rationale for implementing BSC is that users should realise more improvements in their utilisation of non-financial performance measures than non-users. The basis for this rationale can be summarised in the following null hypothesis:

***H.2:** There is no difference between BSC users and non-users in terms of their use of performance measures.*

### **Contingency theory framework**

Otley (1980) notes that the main hypothesis of contingency theory is that there is no commonly relevant accounting system for organisations based in different locations. However, the selection of suitable systems will depend on the particular circumstances of each organisation. Chenhall (2003) argues that contingency studies are mainly concerned with addressing the contingent nature of management accounting practices and with how management accounting practices might be affected by a selection of contingent variables. The main reasons for adopting the contingency theory framework in this paper are threefold. First, contingency theory provides possible explanations for different uses of management accounting practices, including performance measures (Otley, 1980). Second, the contingency framework supplies the best analytical base (Jones, 1985). Finally, as an empirical framework, contingency theory will allow this study to develop relevant hypotheses and to analyse them statistically, depending upon the factors selected (Xiao et al., 1996).

### **Factors influencing performance measures**

Since the 1980s, many studies have focussed on different aspects of management accounting practices, especially performance measurement. These include contextual factors such as increased competition, technological development, and environmental uncertainty, which we will now examine in more detail.

### **The impact of organisational structure**

Laitinen (2006) and Lorenzo (2008) conclude that organisational structure is one of the most important factors affecting management accounting practices. Top management may make some modifications in their organisational structure to become more effective and efficient, and to gain a bigger market share and assure the survival of the firm (Hoque, 2005). Bititci et al. (2002) found that organisational structure should be interdependent throughout the management lifecycle with accounting practices like performance measurement. Successful implementation of these practices will lead to a more participative and consultative management style and may result in significant performance improvements. However, Cobb et al. (1995) argue that changes in organisational structure have an indirect effect on management accounting practices because such shifts are accompanied by changed priorities, which may affect management accounting practices. Organisational structure (centralization/decentralization) is considered an important variable influencing the design of management accounting systems. In addition, the issue of authority and power distribution is crucial to an understanding of the control processes within an organisation (Waterhouse and Tjessen 1978).

Gordon and Miller (2003) note that as the administrative task becomes more complex, subtasks and responsibilities must be delegated to lower levels of management to ease the burden of decision making. Thus increased environmental dynamism, heterogeneity, and hostility are often accompanied by decentralisation of power and responsibilities. They add that under these conditions, accounting information systems may have to become more sensitive and sophisticated. This may mean, for example, replacing informal controls and increasing the requirements for formal controls, and producing more explicit reports on the performance of organisational subunits (i.e., the accounting system itself becomes decentralised). Hence the third hypothesis is:

***H.3:** Libyan banks that are more decentralised tend to use non-financial measures.*

### **The impact of competition**

There is empirical evidence indicating that business organisations desire appropriate management accounting practices to meet growing competition (Johnson and Kaplan, 1987; Laitinen, 2006). Cooper and Ezzamel (2004) and O'Connor et al. (2008) argue that the new competitive environment is characterised by the motivation for change in every aspect of the organisation, including management accounting practices such as the costing system and the performance measurement system. In addition, Hoque (2005) and Abdl-Maksoud (2008) conclude that competition is one of the important reasons why organisations employ specific practices such as non-financial performance measures. Regarding the effect of competition on

performance measurement systems, a number of studies contend that conventional performance measures are inappropriate given today's complex competitive environment. As a result, much attention has been given to the need for multiple performance measures (Neely et al., 2001; Hussain, 2005; O'Connor et al., 2008). As a result, there has been much concern about the need for multidimensional performance measures (Kaplan and Norton, 1996; Neely et al., 2001; Hussain and Gunasekaran, 2002). In view of the stiff competition faced by Libyan banks, this paper selects the effect of competition for investigation and sets forth the following hypothesis:

**H.4:** *Libyan banks that are facing competition tend to use financial and non-financial measures for performance evaluation.*

### **The impact of organisation size**

In response to such economic pressures, management accounting practices adapt to their environment with various degrees of responsiveness. However, characteristics such as the size and type of a company are key to determining its potential for change and adaptation to economic pressures (Granlund and Lukka, 1998; Hussain and Gunasekaran, 2002). As for as the impact of organisation size on performance measurement systems, several previous studies (Chenhall, 2003; Ezzamel, 1990) suggest that top management in large firms will implement a multiplicity of performance measures (relative to small firms) to motivate managers of different responsibility centres. For example, Chenhall (2003) indicates that size indeed affects the design of performance measurement systems: larger organisations use more sophisticated performance evaluation systems and tend to introduce non-financial measures. In addition, organisation size might influence the shape of the control systems, which tend to be more sophisticated within bigger firms (Libby and Waterhouse, 1996; Speckbacher et al., 2003). This study argues that in the Libyan banking sector, organisation size may influence the use of financial and non-financial measures. Previous research has indicated that size indeed affects the design of performance measurement systems. Larger organisations employ more sophisticated versions of such systems (Chenhall, 2003) and tend to introduce non-financial measures (Hoque and James 2000). This results in the following hypothesis:

**H.5:** *Organisation size is positively associated with the use of financial and non-financial performance measures.*

### **The impact of oriented business strategy**

Otley (1980), Chenhall (2003), and Juson et al. (2008) state that performance measurement systems should be designed to work with the business strategy of the organisation, suggesting that the choice of performance measurement systems is dependent on business strategy. In addition, the nature of performance measurement systems differs according to the type of business strategies employed to achieve long-term competitive advantage (Hussain, 2004; Lorenzo, 2008). Stede et al. (2006) study the relationship between business strategy and the use of non-financial performance measures by Belgian and U.S. managers; they found that there is a positive relationship between business strategy and the extent to which non-financial performance measures are used. However, Langfield-Smith (1997) and Verbeeten and Boons (2008) maintain that the influence of business strategy on the performance measurement system is not clear and is in fact limited. The type of strategy adopted should influence the use of performance measures. Miles and Snow (1978) described patterns of behaviour used by organisations to adjust to their environments. They identified four strategic types of organisation: prospectors, defenders, analysers, and reactors.<sup>1</sup>They confirmed that

---

<sup>1</sup>Prospectors are mainly entrepreneurial and oriented towards innovation and new opportunities. Using a variety of means, defenders of the existing market target a narrow (or niche) market sector. Analysers are a hybrid of the prospector and defender types, while reactors are those who act in response to change.

each type has its own unique strategy for relating to its chosen market, and that each has a particular configuration of technology, structure, and process that is consistent with its market strategy. Therefore the last hypothesis is:

**H.6:** *The type of strategy orientation tends to affect the use of financial and non-financial measures.*

## **Libya and its banking environment**

Prior to the discovery of oil, Libya was one of the poorest countries in the world (Wright, 1981; Vandewalle, 1998). There was a deficit in the budget and the balance of payments, and most government projects were likewise operating in arrears (Higgins, 1959). Farley (1971), a United Nations (UN) economist, pointed out that before the discovery of oil the Libyan economy was underdeveloped, with no indication of any economic growth.

The economy became primarily dependent on the revenues from the oil sector from 1961 onwards (Wright, 1981; Vandewalle, 1998). These oil revenues and a small population give Libya one of the highest gross domestic products (GDP) per person in Africa. During the period of 1951 (the year of independence) to 1969 (the Libyan revolution), the Libyan economic system was mainly capitalist.

After 1970 several steps were taken by the revolutionary government to reform the economic situation, and it changed from a capitalist to a socialist economy. State intervention in the economy increased and the government started expanding the public sector and reducing the private sector. From the late 1970s to 1991, the Libyan economy was centrally planned and the state had managerial control of all manufacturing activities, foreign and domestic retail trade, banking, and insurance services. However, private companies started to operate in Libya from the 1990s onwards. Moreover, the Libyan national economy faced technical sanctions (imposed by the United States and the U.N.), which hindered industrial enterprises from importing technology. As a result, of these crises, the state introduced a series of economic and political liberalisation measures, which permitted the private sector to play a significant role.

In 1992, the state issued Act No. 9 to enhance private sector activities in the national economy to enhance economic development and open the door for the privatisation of a number of public sector companies. Since the early 2000s, the state has continued its economic reforms as it attempts to rejoin the international community. It has applied for membership of the World Trade Organisation (WTO), and many state-run industries are being privatised. U.S. and U.N. sanctions were lifted and many international oil companies have returned to the country.

The banking sector is one of the most important in Libya. It is also one of the most sensitive structural components of any economy. Historical sources show that the first banks in Libya were established in the beginning of the nineteenth century, at the end of the Ottoman period. Until the middle of the twentieth century, the commercial banks were branches of foreign banks, such as Barclays (Al-Arbah, 1985).

The Libyan banking sector now consists of the Central Bank of Libya (CBL), specialised banks (the Libyan Arab Foreign Bank, the Agricultural Bank, the Savings and Investment Bank, and the Development Bank) and commercial banks. The commercial banks have an economic and social role. Libyan commercial banks consist of public commercial banks (state commercial banks [SCBs]) and private commercial banks (Central Bank of Libya, 2001). They played a significant role in the growth of the Libyan economy.

The period of 1972–1992 was dominated by five public commercial banks. On 13 November 1969 a decree was issued that required Libyanise foreign banks to become Libyan joint stock companies. Libyan nationals were obliged to own more than 51% of the company and the majority of its board of directors, including the chairman, had to be Libyans. In December 1970, Law No. 153 nationalised foreign shares in the commercial banks and specified the contribution of Libyans to them. It also reorganised the banks and increased the contribution of the Central Bank of Libya to 51% in banks where it was less than that. In order to keep up

with the latest developments in both the national and international environment, legislation was enacted in the 1990s to encourage the private sector to participate in owning and managing commercial banks (Masoud and Al-Shrif, 2002).

From 1993 onwards the banking sector witnessed important developments. Law No. 1 1993 (amended by Law No. 1 2005) allowed private banks to be established and permitted foreign banks to open branches, agencies, or representative offices. The private banks established following this passage of this law include the Bank of Commerce and Development, Aman Bank for Commerce and Investment, Al-Ijmaa Al-Arab Bank, Al-Wafa Bank, Representative Office of Jordanian Housing Bank, and 48 small private banks (Luxford, 2005 and Central Bank of Libya, 2004). These banks now compete very effectively with state banks. To observe international standards, the CBL has issued regulations for the commercial banks. Those handed down by Basle Committee, for example, are aimed at assuring that capital investment is in line with international developments and innovations, and at reaching a banking standard necessary to compete in the international banking world.

### **Research method and survey instrument**

Our research was carried utilising a questionnaire preceded by an introductory letter clarifying the purposes and objectives of the entire project. Our sample consists of more than 63 respondents from the Libyan banking sector, including commercial and specialised Libyan banks. Among this group, the researchers contacted the branch managers directly in order to select a list of banks prepared to cooperate with the research. The survey was conducted by sending a questionnaire during the second half of 2008. After three follow-up phone calls made to non-respondents to increase survey response rate, 93 questionnaires (63 usable) were sent back. The final response rate of about 67% represents an acceptable target when the questionnaire involves top and middle management (senior managers and branch managers). The questionnaire was developed and refined as follows: nearly all items in the performance measures and contingent factors categories were adapted from previously published works. A preliminary draft of the questionnaire was discussed with the supervision team and some research students at Liverpool John Moores University to assess the content validity prior to pilot testing. A pilot test was conducted with a group of five branches, whose inputs were used to improve the clarity, comprehensiveness, and relevance of the survey instrument.

The questionnaire was structured in two parts. In the first part, organisations were asked to indicate on a five-point Likert scale of importance—where 1 = not at all, 3 = moderately, and 5 = extensively—the extent to which they used a set of performance measures derived from academic/practitioner management accounting literature (White 1996; Kaplan and Norton 1996, 2000; Gosselin 2005). The second part listed some contingency factors such as decentralisation, strategic types of organisation, and innovative management accounting techniques (e.g., BSC).

### **Sample features**

Our data was analysed using the SPSS package v15.0. The reliability of the questionnaire was also verified. Internal consistency was established using Cronbach's Alpha; it was equal to .820. The first set of empirical results was produced through the use of descriptive statistics. Table 1 indicates information such as the date the business was established, its type, ownership, total assets, and type of business strategy. Tables 2, 3, and 4 describe the distribution of respondents according to the importance placed on bank success and the extent of current use of performance measures. In addition, the measures were adjusted according to ownership and the type of business conducted by the bank.

Table 5 indicates that BSC is systematically applied by 44% and implemented by 16% of the sample. However, about 40% of Libyan banks never considered implementing BSC. Their adoption does not seem to be related to organisation size, as confirmed by Hoque and James (2000) for the BSC case. To test this empirically an Independent Samples t-Test was undertaken. It failed to detect any significant difference between large organisations and smaller ones (Table 6).

Organisational structure (Table 7) is a construct measured by four items on a five-point Likert scale (1 = low, 3 = moderate, and 5 = high). It indicates the degree of authority/delegation in

the organisation. Table 8 notes that the level of competition among Libyan banks is of moderate intensity and that it originates principally in private banks.

### **Findings and hypothesis testing**

To test H.1, the financial and non-financial performance measures were ranked according to the mean of the extent to which respondents from Libyan banks ranked them as important to long-term success and used them in the aforementioned practises. Table 9 incorporates Tables 3, 4, and 5 and accounts for the overall diverse measurements. The last column highlights this indicator, which was calculated by average standardised rating of importance and uses for each category (financial and non-financial measures). This indicator shows that if the level of overall diverse measurements is up to 3, then banks use diverse sets of performance measures at a high level. However, if the rate is less than that, then there is not a high level of use of diverse sets of performance measures. Thus, Table 9 clearly indicates that Libyan banks are still relying on financial performance measures.

The highest rate of overall diverse measurement is financial measures as ranked by mean (3.530). Other non-financial measures ranked less than the level of absenteeism (+3). Therefore H.1 is confirmed.

### **Contingency factors and performance measures**

Theoretically BSC implementation should improve the adoption and the use of non-financial performance measures (customer, employee, quality, and community) as compared to banks that do not adopt BSC. To test empirically H2 an Independent Samples t-Test was used. The results reveal that BSC implementers use performance measures, especially non-financial ones, to a greater extent than other banks. Table 10 includes the means of performance measures of two groups where BSC was either applied or not applied. The table shows the measures, ranked by mean, for which there is a statistically significant difference in the means. It indicates that banks that applied BSC use non-financial measures to a greater extent than other banks, even if the distinction is not completely clear-cut. Further analyses of the data are necessary to better determine whether BSC adoption enhances the use of non-financial performance measures. But at this point H2 seems to be confirmed.

A factor analysis was undertaken in order to classify the measures into categories and to determine the underlying themes among the eight items. Principal Component Analysis (PCA) (Table 11) reveals two interpretable factors with Eigen values greater than 1 that account for 64% of the variance. The two factors are labelled as follows: competition (four items) and decentralised (four items).

To test the remaining four hypotheses a bivariate correlation was undertaken between the four factors, two of them coming from the PCA (competition and decentralisation), type of business strategy (Miles and Snow's [1978] organisation strategies, namely, defender, prospector, and analyser), and size of organisation (total assets).

Table 12 shows all the results of this analysis. Kendall's tau (t) association coefficients help to determine whether there are some associations among four factors. These estimates are accompanied by p-values from statistical significance tests. Decentralisation is positively correlated with quality, financial, employee, and customer measures, respectively, but less correlated with community measures. These results corroborate the idea that organisations that are more decentralised tend to use more non-financial measures. Hence H.3 is confirmed.

The level of competition is positively correlated with financial and non-financial performance measures (even if these values are not statistically significant). So H.4 (Libyan banks that are facing competition tend to use financial and non-financial measures for performance measurement) is not confirmed.

Bank size is positively correlated with both non-financial and financial performance measures; however, with respect to the latter, the correlation is not significant. Thus, H.5 (organisation size is positively associated with the use of financial and non-financial performance measures) is confirmed.

Prospectors are positively correlated with all performance measures, while defenders are negatively correlated with non-financial performance measures (even if this value is statistically significant). Furthermore, analysers are positively correlated with non-financial



performance measures, but this is not the case with financial measures. Overall, these results appear consistent with Miles and Snow's (1978) theory, since prospectors are positively correlated with use of non-financial performance measures, while defenders are negatively correlated.

Therefore, H.6 (the type of strategy orientation tends to affect the use of financial and non-financial measures) is confirmed.

## **Discussion**

These findings are consistent with Miles and Snow's (1978) strategic type of organisations theory, given that prospector organisations "value being first-in" in new products and market areas, even if not all of these efforts prove to be highly profitable. These organisations respond rapidly to early signals concerning areas of opportunity, and these responses often lead to a new round of competitive actions (Snow and Hrebiniack, 1980). Hence they should rely more on non-financial measures than defenders.

For their part, defenders will try to protect their domain by offering higher quality, superior service, lower prices, and so forth. Often they are not at the forefront of developments in their industry and tend to ignore industry changes that have no direct influence on current areas of operation and concentrate instead on doing the best job possible in a limited area (Snow and Hrebiniack, 1980). Defenders tend to place stronger emphasis on customer-related measures. In fact, the Kendall coefficient between customers is greater than that between prospectors, even if both are not significant. The level of competition is positively weakly associated with financial and non-financial performance measures, but it is not significant. That does not confirm the hypothesis that when banks are in a turbulent or unstable environment, they tend to rely mostly on financial measures (Gosselin, 2005). BSC is positively associated with bank size. That represents a further confirmation of H.2. Banks implementing BSC use more non-financial performance measures.

## **Conclusions and implications**

Contingency theories have influenced much of the empirical work in the management accounting field, emphasising, especially, the role of both environment and strategy. The latter represent complex problems for companies, as do other contingency factors, such as dimension, technology, organisation design, and management accounting systems.

This exploratory study has been designed to test some specific contingency relationships between competition, decentralisation, strategic type of organisation, innovative management accounting techniques (e.g., BSC) with the adoption of performance measures. Literature on management accounting has investigated the relationships from a contingency viewpoint (Jones, 1985; Chenhall and Morris, 1986). It has examined, for instance, the relationship between business strategy and management control systems (Otley 1980), on the conviction that management accounting systems can be identified as an organisational variable whose consistency (structure, shape, characteristics, and composition) will depend on a series of circumstances that firms will be forced to face during their existence.

While Skinner (1969) concludes that the relationship between operations and corporate strategy is not easily understood, the focus has recently shifted towards empirical research on performance measurement systems. This suggests that firms may use types of measures that fit with their strategies (Gosselin 2005). For example, environmental uncertainty (Hoque 2005), advanced management techniques and advanced technologies (Maksoud et al., 2005) are used in accordance with contingency theory on management accounting.

The main purpose of this paper was to collect empirical evidence on a set of performance measurements implemented in Libyan banks. Despite literature suggesting firms should increase the adoption of non-financial measures in their performance measurement systems, the results of this research confirm that financial measures continue to be used more frequently by managers and controllers.

A further purpose was to evaluate the extent to which top managers and controllers used BSC. The systematic application of BSC methodology seems to be low. Ittner and Larcker (2003) report that most companies have apparently adopted versions of non-financial measurement frameworks (e.g., Kaplan and Norton's BSC) but they have failed to align cause-and-effect

relationships with firm strategy. This suggests that managers may be manipulating performance measures in order to raise earnings and bonuses. Hence, paradoxically, it appears that non-financial measures are just as, if not more, susceptible to manipulation than financial accounting measures.

This brief paper gives an initial account of the application of performance measures in Libyan banks, explaining their use in a BSC setting. Overall, the results of the research confirm that Libyan bank managers are still relying on financial performance measures, even if less so when BSC is applied. Financial accounting data is useful, but probably more so if it is integrated with non-financial performance measures. Specifically, our findings confirm a positive trend among managers dealing with BSC: they are attempting to avoid an unbalanced emphasis on relating incentives to quantifiable financial statement measures (i.e., ROI, ROS, EBITDA, and Gross Margin). Indeed, using only incentives that are tied to short-term performance measures can lead managers to focus heavily on short-term gains (Kaplan and Norton 1996) rather than on drivers more suitable for long-term firm value (Eccles, 1991).

### Limitations of the study

The aim of this research is to contribute to a better understanding of which performance measures are used by managers. Specifically, this paper upgrades the existing theory; establishes relationships between contingency factors and performance measures with contingency theory; and produces some results that would be interesting to develop further. However, the study has some limitations. First, the sample is derived from Libyan banks and does not consider other sectors (e.g., manufacturing). Furthermore, the paper does not consider how these contingency relationships may affect organisational performance and which combinations of performance measures can lead to improved financial results and organisational behaviour with more regular use.

### Tables

**Table 1:** Description of banks according to the classifications of the survey.

<b>Classification of banks by year of establishing business typology</b>				
<b>Number (%)</b>	<b>Before 1980</b>	<b>Between 1981 and 1990</b>	<b>Between 1991 and 2006</b>	<b>After 2006</b>
	52 (76.5)	0 (0)	16 (23.5)	0 (0)
<b>Classification of banks according to type of business</b>				
<b>Number (%)</b>			<b>Commercial</b>	<b>Specialized</b>
			51 (75.0)	17 (25.0)
<b>Classification of banks by ownership typology</b>				
<b>Number (%)</b>	<b>State-owned (public bank)***</b>	<b>Private</b>		
		<b>PAPP*</b>	<b>PSE**</b>	
	35 (51.5)	17 (25.0)	16 (23.5)	
<b>Classification of banks by total of assets typology</b>				
<b>Number (%)</b>		<b>Less than 100</b>	<b>100-500</b>	<b>Above 500</b>
		11 (16.5)	27 (39.7)	30 (44.1)
<b>Classification of banks by type of strategy typology</b>				
<b>Number (%)</b>		<b>Prospector</b>	<b>Analysers</b>	<b>Defender</b>
		28 (41.2)	18 (26.5)	22 (32.4)

\***PAPP** = private after process of privatisation (before that they were public).  
 \*\* **PSE** = private since establishing. \*\*\***State-owned public bank** = the state owns more than 50% of the shares.

**Table 2:** Importance of performance measures to bank success.

	Mea n	Std. D.	Comparisons of survey results by typologies			
			Ownership		Business	
			S-O	P	C	SP
<b>Financial</b>	3.70	0.77	3.60	3.814	3.68	3.76
<b>Customer</b>	2.98	0.83	2.85	3.116	3.02	2.88
<b>Employee</b>	2.32	0.74	2.25	2.388	2.41	2.05
<b>Quality</b>	3.20	0.90	3.17	3.234	3.19	3.23
<b>Community</b>	2.25	0.65	2.28	2.217	2.25	2.23

UNL = unlisted banks. L = listed banks. S-O = public banks. P = private banks. C = commercial banks. SP =

**Table 3:** Use of performance measures for managerial performance evaluation.

	Mea n	Std. D.	Comparisons of survey results by typologies			
			Ownership		Business	
			S-O	P	C	SP
<b>Financial</b>	<b>3.8</b>	0.85	3.60	4.031	3.90	3.52
<b>Customer</b>	<b>2.7</b>	0.84	2.42	3.028	2.90	2.17
<b>Employee</b>	<b>2.1</b>	0.69	1.74	2.487	2.29	1.52
<b>Quality</b>	<b>2.7</b>	1.22	1.82	3.756	3.13	1.64
<b>Community</b>	<b>1.8</b>	0.76	1.51	2.184	2.02	1.29

**Table 4:** Use of performance measures to identify problems.

	Mea n	Std. D.	Comparisons of survey results by typologies			
			Ownership		Business	
			S-O	P	C	SP
<b>Financial</b>	<b>3.07</b>	0.997	2.48	3.691	3.31	2.35
<b>Customer</b>	<b>2.26</b>	0.803	1.77	2.789	2.45	1.70
<b>Employee</b>	<b>1.79</b>	0.612	1.57	2.026	1.90	1.47
<b>Quality</b>	<b>2.51</b>	0.985	1.74	3.325	2.82	1.58
<b>Community</b>	<b>1.66</b>	0.563	1.60	1.728	1.68	1.58

**Table 5:** Rate of adoption of BSC technique

	Extent considered and implemented			Mean	Std. D.	Comparisons of survey results by typologies			
	1	2	3			Ownership		Business	
						S-O	P	C	SP
<b>Balanced scorecard</b>	39.7	44.1	16.2	<b>1.765</b>	0.715	1.229	1.958	1.941	1.235
1 = never considered. 2 = considered for implementation. 3 = implemented.									

**Table 6:** BSC and bank size (total assets)

	LB	SB	Mean LB	Mean SB	Mean difference	Std. error difference	Sig.
<b>Rate of BSC implementation</b>	<b>30</b>	<b>38</b>	<b>1.2667</b>	<b>2.2728</b>	<b>1.0061</b>	<b>.44978</b>	<b>.408</b>
<b>LB=</b> Large Banks which total assets less than 500 Million, <b>SB=</b> Smaller Banks which total of assets is more than 500 Million							

**Table 7:** Organisational structure.

Degree of decentralisation in the bank	Level of use			Mean	Std. D.	Comparisons of survey results by typologies			
	1 / 2	3	4 / 5			Ownership		Business	
						S-O	P	C	SP
<b>Delegation of authority</b>	70.6	26.5	2.9	2.118	0.763	1.657	2.608	2.314	1.529
<b>Feedback to employees</b>	76.4	23.5	0	1.897	0.756	1.371	2.454	2.098	1.294
<b>Formalisation and job description</b>	47.1	35.3	17.6	2.588	1.011	2.000	3.221	2.824	1.882
<b>Level of decision making taken at the top level of management</b>	1.5	51.5	47.1	3.603	0.756	3.657	3.550	3.706	3.294

**Table 8:** Market competition and origin of competition.

	Level of intensity of competition			Mean	Std. D.	Comparisons of survey results by typologies			
	1	2/3	4/5			Ownership		Business	
						S-O	P	C	SP
<b>Competition on prices</b>	17.6	78	4.4	2.236	0.794	2.200	2.272	2.353	1.882
<b>Competition on quality and variety of service</b>	29.4	69.1	1.5	1.985	0.782	1.943	2.033	1.980	2.000
<b>Competition on gaining bigger market share</b>	47.1	53	0.0	1.588	0.604	1.457	1.721	1.667	1.353
<b>Competition relating to customers</b>	22.1	69.1	8.8	2.206	0.890	1.771	2.662	2.392	1.647
<b>Origin of competition number (%)</b>	<b>Public banks</b>			<b>Private banks</b>			<b>Foreign banks</b>		
	16 (23.5)			42 (61.8)			10 (14.7)		

1 = negligible intensity. 2 / 3 = low to moderate intensity. 4/5 = high to extremely high intensity.

**Table 9:** Extent of performance measurements use.

		Extent of importance and use of performance measurements			Overall diverse measurements
		Importance of performance measures	Use for managerial evaluation	Use for identifying problems	
<b>Financial</b>		3.706	3.809	3.074	<b>3.530</b>
<b>Non-financial</b>	<b>Customer</b>	2.985	2.721	2.265	<b>2.657</b>
	<b>Employee</b>	2.324	2.103	1.794	<b>2.074</b>
	<b>Quality</b>	3.206	2.765	2.515	<b>2.829</b>
	<b>Community</b>	2.250	1.838	1.662	<b>1.917</b>

1 = rarely used. 3 = moderately used. 5 = frequently used.

**Table 10:** Performance measures with balanced scorecard.

Performance measures	Number of banks adopting BSC	Number of banks not adopting BSC	Mean BSC applied	Mean BSC not applied	Mean difference	Sig.
<b>Financial</b>	11	27	4.3636	3.5926	0.771	0.223
<b>Customer</b>	11	27	3.2727	2.5556	0.7171	0.263
<b>Employee</b>	11	27	2.5455	1.8148	0.7307	0.47
<b>Quality</b>	11	27	4.000	1.8889	2.1111	0.912
<b>Community</b>	11	27	2.000	1.5556	0.4444	0.063

**Table 11:** Factor analysis: rotated factor matrix for eight items.

	Components	
	Competitio n	Decentralis ed
<b>Competition on prices</b>	0.5778	
<b>Competition on quality and variety of service</b>	0.6464	
<b>Competition on gaining bigger market share</b>	0.8329	
<b>Competition relating to customers</b>	0.8651	
<b>Delegation of authority</b>		0.6326
<b>Feedback to employees</b>		0.7431
<b>Formalisation and job description</b>		0.5878
<b>Level of decision making taken at the top level of management</b>		0.5491
<b>Eigen values</b>	4.231	4.148
<b>% of variance</b>	10.07	9.875
<b>Cumulative %</b>	10.07	19.95

**Table 12:** Correlation matrix tau (t) Kendall association measure.

	Size of bank	Competition	Decentralised	Analysers	Defender	Prospector
<b>Financial</b>	0.079	0.065	.334(**)	0.138	.211(*)	.298(**)
<b>Customer</b>	.192(*)	0.122	.282(**)	.358(**)	-	.376(**)
<b>Employee</b>	.432(**)	0.001	.330(**)	.454(**)	-	.478(**)
<b>Quality</b>	.578(**)	0.032	.370(**)	.498(**)	-	.719(**)
<b>Community</b>	.407(**)	0.133	.210(*)	.403(**)	.406(**)	.335(**)
*Correlation is significant at the 0.05 level (two-tailed).						
** Correlation is significant at the 0.01 level (two-tailed).						

### Acknowledgements

I am thankful to a number of PhD students at Faculty of Business and Law at Liverpool John Moores University for their substantial cooperation during the research project. I am also grateful for the financial support provided by the Libyan Bureau Cultural Affairs in London.

## References

- Abdel-Maksoud, A., Asada, T., and Nakagawa, M. (2008) 'Performance measures, managerial practices and manufacturing technologies in Japanese manufacturing firms: state-of-the-art,' *International Journal of Business Performance Management* 10 (1), pp. 1–16.
- Abdel-Maksoud A., Dugdale D., and Luther R. (2005) 'Non-financial performance measurement in manufacturing companies,' *The British Accounting Review* 37 (3). pp. 261–297.
- Al-Arbah, S.A. (1985) *The Banking System in Socialist People's Libyan Arab Jamahiriya*. Beirut: The Institute of Banking Business and Management Publishing, Al-Dar Arabia for Cyclopaedia.
- Al-Enzi, A.M.N. et al. (2006) 'Non-financial performance measurement in the banking sector: four grounded theory case studies,' *International Journal of Accounting, Auditing and Performance Evaluation* 3 (3), pp 362–385.
- Anderson, S.W. and Lanen, W. (1999) 'Economic transition, strategy and the evolution of management accounting practices: the case of India,' *Accounting, Organizations and Society* 24 (5-6), pp. 379–412.
- Anthony, R.N. and Govindarajan, V. (2001) *Management Control Systems, 10<sup>th</sup> Edition*. New York: McGraw-Hill.
- Atkinson, A. et al. (1997) 'New directions in management accounting research,' *Journal of Management Accounting Research* 9, pp. 79–109.
- Banker, R.D., Potter, G. and Srinivasan, D. (2000) 'An empirical investigation of an incentive plan that includes non-financial performance measures.' *The Accounting Review* 75 (1), pp. 65–92.
- Berry, R.H., Crum, R. E., and Waring, A. (1993) *Corporate Performance Evaluation in Banking Lending Decisions*. London: CIMA.
- Bititci, U.S. et al. (2002) 'Web enabled performance measurement systems,' *International Journal of Operations and Production Management* 22 (11), pp.1273–1287.
- Bourne, M. and Neely, A. (2002) 'The success and failure of performance measurement initiatives: perceptions of participating managers,' *International Journal of Operations and Production Management* 22 (11), pp. 1288–1310.
- Central Bank of Libya (2001) *Economic Bulletin* 41, First Quarter.
- Central Bank of Libya (2004) *Annual Report No. 47 (2003)*. Tripoli Musrath Al-Ezdihar Printing.
- Chenhall, R.H. (2003) 'Management control systems design within its organizational context: findings from contingency-based research and directions for the future,' *Accounting, Organizations and Society* 28, pp.127–168.
- Chenhall, R.H. and Morris, D. (1986) 'The impact of structure, environment, and interdependence on the perceived usefulness of management accounting systems,' *The Accounting Review* LX1 (1), pp. 16–35.
- Cobb, I., Helliard, C., and Innes, J. (1995) 'Management accounting change in a bank,' *Management Accounting Research* 6, pp. 155–175.
- Cooper, D. and Ezzamel, M. (2004) 'A social analysis of the assumptions of the balanced scorecard: towards a more integrated strategic performance measurement system,' *The 27<sup>th</sup> Annual Congress of the European Accounting Association*. Prague.
- Drury, C. (2004) *Management and Cost Accounting, 6<sup>th</sup> Edition*. London: Thomson.
- Eccles, R.G. (1991) 'The performance management manifesto,' *Harvard Business Review*, March–April.
- Ezzamel, L. (1990) 'The impact of environmental uncertainty, managerial autonomy and size on budget characteristics,' *Management Accounting Research* 1 (2), pp. 181–197.
- Farley, R. (1971) *Planning for Development in Libya: The Exceptional Economy in the Developing World*. New York: Praeger Publishers.
- Fisher, J. (1995) 'Use of non-financial performance measures,' in S.M. Young, ed., *Readings in Management Accounting*. Englewood Cliffs, NJ: Prentice Hall.
- Fitzgerald, L. et al. (1991). *Performance Measurement in Service Business*. London: Chartered Institute of Management Accountants (CIMA).

- Gordon, L. and Miller, D. (1976) 'A contingency framework for the design of accounting information systems,' *Accounting, Organizations and Society* 1 (1), pp. 59–69.
- Gosselin, M. (2005) 'An empirical study of performance measurement in manufacturing firms,' *International Journal of Productivity and Performance Management* 54 (5/6), pp. 419–437.
- Granlund, M. and Lukka, K.S. (1998) 'It's a small world of management accounting practices,' *Journal of Management Accounting Research* 10, pp. 153–179.
- Guenther, T. and Gruening, M. (2002) 'Measuring performance: a cross-sectional empirical survey,' *25<sup>th</sup> Annual Congress of the European Accounting Association*. Copenhagen Denmark.
- Higgins, B. (1959) *Economic Development*. New York: W.W. Norton & Co.
- Hoque, Z. (2005) 'Linking environmental uncertainty to non-financial performance measures and performance: a research note,' *The British Accounting Review* 37, pp. 471–481.
- Hoque, Z. and James, W. (2000) 'Linking balanced scorecard measures to size and market factors: impact on organizational performance,' *Journal of Management Accounting Research* 12, pp. 1–17.
- Hussain, M.M. (2004) 'Organisational strategic orientation and its impact on non-financial performance measurement in the financial services industry,' *Management Research News* 27 (11/12), pp. 115–133.
- Hussain, M.M. (2005) 'Management accounting performance measurement systems in Swedish banks,' *European Business Review* 17, pp. 566–589.
- Hussain, M. and Gunasekaran, A. (2002) 'An institutional perspective of non-financial management accounting measures: a review of the financial services industry,' *Managerial Auditing Journal* 17, pp. 518–536.
- Hussain, M. et al. (2002) 'Accounting standards and practices of financial institutions in GCC countries,' *Managerial Auditing Journal* 17 (7), pp. 350–362.
- Hussain, M. and Hoque, Z. (2002) 'Understanding non-financial performance measurement practices in Japanese banks: a new institutional sociology perspective,' *Accounting, Auditing and Accountability Journal* 15 (2), pp 162–183.
- Ittner, C.D. and Larcker, D.F. (1998) 'Innovations in performance measurement: trends and research implications,' *Journal of Management Accounting Research* 10, pp. 205–238.
- Ittner, C.D. and Larcker, D.F. (2003) 'Coming up short on non-financial performance measurement,' *Harvard Business Review* 81 (11), pp.88–95.
- Ittner, C.D., Larcker, D.F., and TAYLOR, R. (2003) 'Performance implications of strategic performance measurement in financial services firms,' *Accounting, Organizations and Society* 28 (7/8), pp. 715–741.
- Johnson, H. and Kaplan, R. (1987) *Relevance Lost: The rise and Fall of Management Accounting*. Boston: Harvard Business School Press.
- Jones, C.S. (1985) 'An empirical study of the evidence for contingency theories of management accounting systems in conditions of rapid change.' *Accounting Organizations and Society* 10 (3), pp.303–328.
- Juso, R., Ibrahim, D., and Zainuddin, Y. (2008) 'Business strategy-balanced scorecard measures alignment: an empirical test of its performance implications using systems approach,' *Journal of Global Business Advancement* 1 (2/3), pp.252–270.
- Kald, M. and Nilsson, F. (2000) 'Performance measurement at Nordic companies,' *European Management Journal* 18 (1), pp.113–127.
- Kaplan, R. and Atkinson, A. (1998) *Advanced Management Accounting, 3<sup>rd</sup> Edition*. London: Prentice Hall International.
- Kaplan, R. and Norton, D. (1992) 'The balanced scorecard measures that drive performance,' *Harvard Business Review* 70 (1), pp. 71–79.
- Kaplan, R. and Norton, D. (1996) 'Using the balanced scorecard as a strategic management system.' *Harvard Business Review* 74 (1), pp. 75–85.
- Kaplan, R. and Norton, D. (2001) 'Transforming the balanced scorecard from performance measurement to strategic management,' *Accounting Horizons* 15 (1), pp. 87–104.



- Laitinen, E.K. (2006) 'Explaining management accounting change: evidence from Finland,' *International Journal of Accounting, Auditing and Performance Evaluation* 3 (2), pp. 252–281.
- Langfield-Smith, K. (1997) 'Management control systems and strategy: a critical review,' *Accounting, Organizations and Society* 22 (2), pp. 207–232.
- Libby, T. and Waterhouse, J.H. (1996) 'Predicting change in management accounting systems,' *Journal of Management Accounting Research* 1, pp.137–150.
- Lorenzo, L. (2008) 'Performance measures from a contingency perspective: field evidence,' *31<sup>st</sup> Annual Congress of the European Accounting Association*. Rotterdam.
- Luxford, K. (2005) 'Libya beats a path to foreign bankers,' *Euromoney* 36 (437), pp. 262–265.
- Masoud, A.M. and Al-Shrif, B. (2002), 'The Libyan commercial banks: the development and activation,' *Journal of Arab Banks Union* 259, pp. 81–88.
- Miles, R.E. and Snow C. (1978) *Organizational Strategy, Structure, and Process*. New York: McGraw-Hill.
- Neely, A. (1999) 'The performance measurement revolution: why now and what next,' *International Journal of Operations and Production Management* 19 (2), pp. 205–228.
- Neely, A., Adams, C., and CROWE, P. (2001) 'The performance prism in practice,' *Measuring Business Excellence* 5 (2), pp. 6–12.
- O'Connor, N., Sandra, V., and Francis, C. (2008) 'Performance consequences of the "fit" between management accounting and control systems and the environment,' *31<sup>st</sup> Annual Congress of the European Accounting Association*. Rotterdam.
- Ostinelli, C. and Toscano, G. (1994) 'Putting quality to work in banking through management accounting systems: three Italian banks,' *17<sup>th</sup> Annual Congress of the European Accounting Association*. Venice.
- Otley, D. (1980) 'The contingency theory of management accounting: achievement and prognosis,' *Accounting, Organizations and Society* 5 (4), pp.13–28.
- Otley, D. (1999) 'Performance management: a framework for management control systems research,' *Management Accounting Research* 10 (4), pp. 363–382.
- Skineer, W. 'Manufacturing missing link in corporate strategy,' *Harvard Business Review*, 47, pp. 136–145.
- Snow, C.C., Hrebiniak, L.G. (1980) 'Strategy, Distinctive Competence, and Organizational Performance,' *Administrative Science Quarterly* 25 (2), pp. 317–336.
- Speckbacher, G., Bischof, J., and Pfeiffer, T. (2003) 'A descriptive analysis on the implementation of balanced scorecards in German-speaking countries.' *Management Accounting Research* 14 (4), pp. 361–387.
- Vandewalle, D. (1998) *Libya since Independence: Oil and State-Building*. Ithaca: Cornell University Press.
- Verbeeten, F.H.M. and Boons, A.N.A.M. (2008) 'Strategic priorities, performance measures and performance: an empirical analysis in Dutch firms,' *European Management Journal* 27 (2), pp. 113–128.
- Waterhosue J.H. and Tiessen, P. (1978) 'The contingency theory of managerial accounting: a comment,' *The Accounting Review* April, pp. 523–529.
- Wright, J. (1981) *Libya: A Modern History*. London: Croom Helm.
- Xiao, Z., Dyson, J. R., and POWELL, P.L. (1996) 'The impact of information technology on corporate financial reporting: a contingency perspective,' *British Accounting Review* 28, pp.203–227.

# Appraisal for NHS doctors: a political wand or window dressing? (working paper)

**Cao Fang, Frank Worthington & Elaine Eades**, *The Management School, University of Liverpool*

## Abstract

The issue of performance appraisal (eventually linked to revalidation) within the UK health system has prompted considerable debate and discussion since it was introduced in 2002, as a compulsory process, for individual doctors to demonstrate their fitness to practice. This paper examines the introduction of policies designed to more closely regulate the medical profession, concentrating in particular on the role that the recently introduced system of appraisal leading to revalidation is intended to play in this process. The key implication of the empirical findings is that appraisal is less likely to be effective as a control mechanism for doctors where self-regulation has long dominated. Therefore, if the appraisal process eventually links to revalidation, more attention should be paid to the dynamics between appraisal and medical self-regulation.

## 1. Introduction

Much of the impetus for the introduction of appraisal has been driven by a series of high profile failures of medical self-regulation within the UK National Health Service (NHS). The complex factors precipitating the need for change in healthcare by government have been the subject of much research in recent years (West, 1988; Strong & Robinson, 1990; Barach & Small, 2000a). Health care reforms have moved from the enforcement of managerial functions within the NHS management system (DoH, 1983), followed by a focus on the efficient use of public funds through market mechanisms (DH, 1989), to dedicated health quality improvement initiatives through clinical governance (DoH, 1998). Many of these government initiatives have implicitly sought different means for greater control and power over the medical profession. The continuous policy shifts have led to "hybrid organizations" in which "hierarchy is strengthened and yet markets and competition are valorized" (Cooke, 2006: 224).

In mid-term the Labour government switched the emphasis to tighter controls and the monitoring of individual performance (DoH, 1999; DoH, 2000; Donaldson *et al.*, 2000), culminating in the drive for appraisal linked to revalidation (henceforth ALR). A spate of public inquiries into medical malpractice have demonstrated "failures in self-regulation and weak links between state and professional regulatory systems" (Kuhlmann & Allsop, 2008: 179). These prompted the policy shift to challenge professional self-regulation and impose a legal duty on doctors for patient safety. Most recently the shift has been to the formal appraisal of doctors linked to revalidation, in which the government's power to grant licenses also extends to the development and surveillance of standards for medical practice. This is now seen as the key initiative changing the NHS (DoH, 2007).

This is a study of the policy of improving health service delivery by the control and enforcement of regulating medical professionals in the UK NHS, aiming to explore the problematic issues around doctors' responses to the current appraisal process. The research contributes to the existing literatures in two aspects. First, it is from an historical context, reviewing the evolution from 1983 to 2007 of government health care policies and their implications, in conjunction with other key initiatives for improved health care quality, for example, ALR. Second, it identifies the problem of appraisal practice in health care, especially its conflict with self-regulation, as the NHS has imported an appraisal process from the private sector where self-regulation is not the norm, and where re-licensing is seldom an issue. This adds a new dimension to the human resource management literature on appraisal practices.

This paper is organised as follows: Section 2 presents the literature reviews, whilst Section 3 explains the research methodology. Section 4 presents the findings from the data analysis; Section 5 contains the discussion, and Section 6 the conclusion.

## 2. Literature review on the government health policy agenda

The question of the control and accountability of doctors has been a long-standing issue for the State. The section reviews the literature around the changing dynamics of the professional autonomy of doctors in the light of the appraisal discourse. The key reports and policy documents from both Conservative and Labour governments (Griffiths, 1983; DoH, 1989; DoH, 1999; DoH, 2000; Donaldson *et al.*, 2000; and DoH, 2007) trace a 25-year history of change through a lens which views the policies as a constant struggle between doctors, seeking to preserve their autonomy, and governments seeking to control doctors' work to improve service delivery, clinical performance and patient safety. The other driving force for change stems from public concern following many high profile medical errors and adverse events, several of which prompted public inquiries. The government response was to reinforce centralised control, emphasising governance and quality control. Appraisal was introduced in *Supporting Doctors, Protecting Patients* (DoH, 1999), which proposed legislation to control medical performance.

### **Conservative policies**

From 1979, the new Conservative government's key political agendas were to reduce the tax burden and roll back the State. Therefore, "a degree of ideologically-based suspicion toward the NHS and a reduction in commitment to the founding principle of the service" emerged (Greener, 2001: 636). The Griffiths Report (DoH, 1983) signalled the imposition of managerialism, with private sector values of efficiency and cost control (Hunter, 1995; Klein, 2000), responsiveness to consumers, and performance measurement. A general management structure was imposed, clinicians were drafted into resource management, and a more assertive management style was valorised (Flynn, 1992; Lewis, 1998; Farbey, *et al.*, 1999). These reforms shifted the traditional balance of power from medical staff to central government or to managers.

In the struggle between doctors' autonomy and management control, the empowerment of medical staff through delegation of responsibilities coincided with "a tightening of bureaucratic control mechanisms; arbitrary on-off systems of surveillance and control and a management style which confused elements of both high trust and low trust" (Cooke, 2006: 240). The reforms created "a subgroup of managerially minded professionals increasingly detached from their initial professional base," with the new clinical directors taking on "new quasi-managerial roles" (Ferlie, 1998: 9). This tended to increase the professional autonomy of doctors and was perhaps an unintended consequence of the reforms.

Facilitated by a perceived NHS funding crisis, the government moved towards "a further tightening of state control over NHS resources and ultimately clinical discretion" (North, *et al.*, 1999: 408) with *Working for Patients* (DoH, 1989). This created an internal market in which responsibilities for purchasing and providing services were separated (Donaldson & Gray, 1998). These 1991 reforms intended to change the course of the history of the NHS (Harrison, 1991; Klein, 1995; Firelbeck, 1996), whilst committing to the principles of healthcare financed through taxation and free at the point of delivery, but dependent on "spending restraint and political accountability" (Greener, 2004: 668). The implicit aim was to solve the problem of managing public services through marketisation and managerialisation to improve service quality, efficiency and responsiveness (Dent, 2005).

An analysis of these reforms may explain individual professional responses of doctors and organizational behaviours in the NHS. The reinforcement of the management role at the expense of doctors accelerated bureaucratisation, with the number of senior managers rising from 1,240 in 1988 to 20,010 in 1993 (DH, *Statistical Bulletin*, 1994). The reforms emphasised "the measurement and public disclosure of organisational performance, and the use of information as instruments of accountability" (Wood, 2002: 15). This was to presage the role of information within the appraisal process to increase surveillance. The medical profession needed to accept more collective responsibility, with professional accountability "transformed into corporate accountability" (Ong, *et al.*, 1996: 90).

Further, the insertion of economic considerations between patient and doctor with a "move from trust to contract" (Klein, 1995: 314), attempted to transform traditional hierarchical bureaucratic control into "a free, competitive market within the framework of a publicly funded

service" (op. cit.: 301). New Public Management was a direct challenge to medical autonomy and the dominance of doctors in the NHS (Dent, 2005) and by shifting the balance of power to management from doctors, the government had "demonstrated its willingness, and ability to push these changes through without involving the medical profession in the policy-making process" (Day & Klein, 1992: 469). Through the managerial roles of clinical and medical directors, the reforms also shifted intra-professional power in respect of resource allocation and decision-making (Ong, *et al.*, 1996: 90). It was a process of re-stratification within the profession, in order to partially control medical practices (Coburn *et al.*, 1997).

### **Labour policies**

The 1997 Labour government determined that the quality of healthcare should be its central NHS reform. The internal market was to be replaced by "integrated care,' based on partnership and driven by performance" (DoH, 1997: Para: 6.11) as a 'third way' of running the NHS, combining "the best from the market approach of the Conservatives and the hierarchical approach of Old Labour"(Powell, 1999: 353). This

presaged further major managerial and regulatory initiatives designed to improve performance and tighten control of the health professions (Cooke, 2006). The government's "intense ideological promotion of the patient interest has been matched by its production of a range of policies designed to restrict the traditional autonomy of doctors" (Salter, 2003: 932).

These initial Labour reforms added new regulators, especially the National Institute for Clinical Excellence, the Health Care Commission and the National Patient Safety Agency, which were partly constituted for target setting and monitoring professional performance. These external medical governance institutions were to be "integrated with the normal procedures of NHS management," the independence of the medical profession was to be restricted, and "its ability to inhibit policy implementation much reduced" (Salter, 2003: 932). Medical governance seemed to be construed as "a new form of centralisation, more like hierarchy than network" (Harrison & Wood, cited in Hann, 2000: 42), and the government attempted to "persuade the medical profession to trade collective autonomy for individual autonomy" (Klein, 1998: 124).

Two years later, the media exposed cases of medical malpractice which led to criticisms of both professional self-regulation and the NHS procedures which failed to prevent, recognise and deal effectively with poor clinical performance. These criticisms hastened the publication of new proposals in *Supporting Doctors, Protecting Patients* (DoH, 1999), based on *Good Medical Practice* (GMC, 1995). This was a revised doctors' code of practice, the purpose of which was to strengthen procedures for professional self-regulation through revalidation and included the requirement to take responsibility for colleagues' performance. Most importantly for this paper, it set out a structure for compulsory appraisal and assessment of doctors' performance linked to revalidation by the GMC (DoH, 1999; Para: 5.18).

Thus the Government's clinical governance initiative and the embryonic GMC ALR arrangements were moving from conception to implementation. This increasing emphasis on appraisal implied a need to design a new system for medical regulation because of the weakness of current traditional models of professional-led regulation. If the GMC is to continue this role, it needs a fundamental change in its governance and culture (DoH, 1999). Revalidation is part of an increased bureaucratic control being applied to professional self regulation (Zwanenberg , 2004: 686) and has put the profession under great pressure.

In *The NHS Plan* (DoH, 2000), the quality agenda expanded through "the bureaucracy of sponsored consumerism" (Salter, 2003: 932), by emphasising patient empowerment and lay representation on the regulatory bodies, including even the General Medical Council. More, it emphasised the need for stronger regulation of professional standards (DoH, 2000: 88), with better leadership, accountability and appraisal. The new consultant contract made their annual appraisal mandatory (DoH, 2000: 97; Para: 8.21). The emphasis on performance monitoring has encouraged managers, cynical from continuous policy changes, to engage in high risk "game playing" in order "to present their organizations in the best possible light in recognition of the potential link between the success of their career and their organizations' success" (Greener, 2008: 204). These initiatives challenged clinical behaviours by emphasising evidence-based medicine and by placing patient choice before medical expertise (Greener, 2001). They placed further responsibility on doctors to raise activity levels, "which their managers feel is required to meet the ever more stringent targets imposed by central

government" Cunningham (2002: 138) and generated additional demands by the imposition of appraisal.

The Chief Medical Officer's report *An Organisation with a Memory* (Donaldson *et al.*, 2000) was a landmark in raising patient safety awareness, demanding organisational learning from adverse events. The report indicated that cultural dissonance between patients and doctors was now a key political issue to drive NHS change to restore public trust in the medical profession. In particular *The Shipman Inquiry* (Smith 2005), *The Alder Hey Inquiry* (Redfern, 2001) *The Ritchie Inquiry* (Ritchie, 2000), *The Neal Report* (Matthews, 2004) and *The Kerr/Haslam Inquiry* (Pleming, 2005) raised issues about the control of individual medical staff (Baker, 2001; 2002). The inquiry into Bristol Royal Infirmary also highlighted the need to examine the climate of NHS organisational culture, communication and control systems (Kennedy, 2001). The CMO suggested that a fundamental culture change is necessary to ensure that measures are introduced to improve quality of health care and reduce adverse events.

The White Paper *Trust, Assurance and Safety, The Regulation of Health Professionals in the 21st Century* (DoH, 2007) was published alongside the Government's response to the Shipman Inquiry and other similar inquiries which have critiqued the effectiveness of self-regulation. The White Paper set out significant reforms to, and legislation for, the regulation of the health profession. However, "a considerable amount of work remains to be done" (Buckley, 2007: 102), especially around a regulatory framework that can simultaneously deal with rogue doctors and the good practice of the vast majority. It must balance perceptions of regulation working in the interests of patient safety with doctors believing the system to be "fair, just and supportive" (op. cit.: 99), whilst not being controlled by the government or any dominant interests. The framework must integrate the functions of standards, ethics, education, registration and fitness to practice and must "enable early intervention, initially at a local level, if a doctor's practice begins to deteriorate" (ibid.). This last point reflects the concern from the public inquiries about how long malpractice had continued unreported. It emphasised that if appraisal is to be effective, robust and consistent, the clinical governance framework within which it operates must be appropriately designed (DoH, 2007).

This 'modernised' regulatory regime brings specific management challenges, to ensure the maintenance of newly introduced quality systems, to make certain that those being audited do not simply develop coping strategies which give the illusion of complying with performance measurement standards "while retaining as much autonomy as possible" (Nettleton *et al.*, 2008: 334-5). These proposals shift the focus from corporate to clinical action. It is the appraisal process, seen by the government as a central component of revalidation (DoH, 2007: 34), that is of the greatest concern to doctors, since revalidation is a mechanism that allows them to demonstrate that they remain up-to-date and fit to practice.

This section has examined twenty-five years of policy-making which have influenced the changing dynamics of the professional autonomy of doctors. The policies have displayed contradictions between "centralization versus decentralization; empowerment versus tighter control; hierarchy versus competition" (Cooke, 2006: 224). The dynamics of policy agendas, demonstrated in each of these seven initiatives have shown that there are certain ambiguities in the control and monitoring of healthcare systems which have increased the conflicts between the medical profession's professional self-regulation on the one hand, and the government's desire for control through regulation and managerialism on the other.

### **3. Research methods**

The research methodology juxtaposes an extensive review of government health policies from 1983 to 2007. These policies exerted intensifying pressures from the State to increase control over professional autonomy to monitor clinical performance and improve service delivery, with evaluated empirical data from questionnaire surveys and interviews with doctors subjected to the appraisal process. Two cohorts were established for the distribution of the questionnaires, namely Non-Consultant Career Grade (NCCG) and Consultant Career Grade (CCG) doctors within the NHS in England. The researcher intends to investigate the interrelationship between doctors' attitudes towards the control mechanism of appraisal.

The questionnaire for Non-Consultant Grades comprised 38 questions in five sections. Of these eight were open questions. The questionnaire for Consultant Grades was broadly similar but comprised 43 questions in six sections, of which 10 were open questions. One section was devoted to those Consultants who had acted as appraisers. There was no compulsion to return the questionnaires. The data has been analysed in this research based on descriptive statistics to identify variables affecting both career grades' response rates and the acceptability of appraisal process. Sequential modelling began with maximal models in which all putative explanatory variables were fitted simultaneously. Terms associated with non-significant changes in scaled deviance were removed sequentially until the minimal models necessary to explain the data were reached.

Some data from several open-end questions included in the questionnaire were subsequently coded, and the analysis of the data proceeded by detailed scrutiny of the selected transcripts to identify key themes. The data were examined for similarities and differences within each key theme, within which several categories were identified. The relationship between the assumptions that appraisal may be a change agent to improve health service delivery, and an efficient control mechanism to improve clinical performance, was explored. These categories were expected to provide explanations of the data.

Following the prior questionnaire survey, volunteers were sought for subsequent follow-up interviews. Of the 52 interviews, 44 were conducted face-to-face and eight by telephone. There were two sets of data examined from the interviews – Consultant and Non-Consultant Grades. All respondents were asked similar questions in random order in the interview. The interviews for Non-Consultant Grade were based on 12 guiding questions, while the Consultant Grade interviews included an additional three questions for those respondents who were appraisers. The interviews were tape-recorded by the interviewer following informed consent by the interviewee. The quotations from interviewees are quoted verbatim and may not be grammatically correct.

#### 4. Analysis of empirical findings from questionnaires and interviews

##### 4.1 Questionnaire survey

The empirical results from questionnaires revealed seven major themes: management attitudes; appraisal functionality; appraisal's link to revalidation; problems with appraisal implementation; the reliability, validity and relevance of data; process strengths and weaknesses; and behavioural changes in response to appraisal.

##### ***Management attitudes to the implementation of appraisal in the Trusts***

**Table 1:** NCCG who have been appraised in the last year  
Explanation of NCCG who have been appraised in the last year

		<b>Frequenc y</b>	<b>Percent</b>	<b>Valid Percent</b>	<b>Cumulative Percent</b>
<b>Valid</b>	<b>No</b>	<b>478</b>	<b>33.4</b>	<b>33.6</b>	<b>33.6</b>
	<b>Yes</b>	<b>944</b>	<b>65.9</b>	<b>66.4</b>	<b>100.0</b>
	<b>Total</b>	1422	<b>99.3</b>	<b>100.0</b>	
	<b>Not applicable</b>	<b>1</b>	<b>.7</b>		
	<b>Total</b>	<b>1432</b>	<b>100.0</b>		

Table 1 shows that 66.4 % of a total of 1422 staff had been appraised in the last 12 months and there were 33.6% of staff who had not been appraised at the time of the survey. Evidence suggests that some Trusts had failed to conduct appraisals because of inefficient management

as there were “no clear arrangements on who will do appraisal and no one has interest in needs of NCCGs” (*Staff Grade 0914*). The survey also shows that some Trusts seemed to lack the initiative to organise appraisals. One respondent stated, “I have to chase and beg to be appraisal annually” (*Senior Clinical Medical Officer 0724*). Another respondent claimed that there was “poor communication of the process of appraisal and the lack of information from the Trust in preparation for it. I got the information and forms from DoH website through my own initiative and had to request an appraisal” (*Staff Grade 0946*).

**Table 2:** CCG who have been appraised in the last year  
Explanation of CCG who have been appraised in the last year

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	34	8.2	8.2	8.2
	Yes	383	91.8	91.8	100.0
	Total	417	100.0	100.0	

Table 2 shows that 91.8 % of the 417 Consultant Grades had been appraised within the last 12 months. In comparing respondents’ rates from both Grades, apparently Consultant Grades gave a higher questionnaire return rate, and more had been appraised than Non-Consultant Grades.

### ***How medical professionals perceive appraisal in terms of its functionality***

The survey results do not provide clear evidence to show that the appraisal process could be functional as a means of identifying poor performance. One respondent viewed his/her appraisal as “overall, seems like [a] duty, not really an assessment of what I have done or want to achieve” (*Staff Grade 1305*). Further consideration is needed over how appraisal could be functional for monitoring poor performance under the current organisation with its present cultural form and professional control methods. One respondent pointed out that “appraisal has no ‘teeth’ in its current organisation. Would prefer an assessment from Department, more in-house discussion” (*Associate Specialist 1004*).

There were concerns that the current appraisal process is not used for improvement of professional performance in some Trusts. Instead it was “used as an opportunity for criticism” (*Staff Grade 0965*). Evidence also showed that the way the appraisal was conducted made the professional “feel strongly that it is verging on being a paper exercise - more regulatory than about enhancing my performance etc.” The respondent further claimed that “the whole exercise has the potential to become meaningless except to prevent future ‘Shipman’ cases” (*Associate Specialist 0877*).

However, one Consultant commented that the current appraisal processes might not in fact detect problem doctors at all because “an incompetent, psychopathic or criminal doctor could easily perform a ‘good’ appraisal” (*Consultant 184*). Therefore, such an appraisal is unlikely to support revalidation effectively. It is suggested that the appraisal process needs re-adjustment to enhance its function as a change agent to improve NHS performance. One respondent has expressed doubts, “I feel, like Audit, it will soon start losing focus and direction - everybody forced to do it for record purpose, but with no effects or improvement” (*Staff Grade Paediatrician 1048*).

### ***Link to revalidation***

In the Non-Consultant Career Grade questionnaires, there were no significant concerns about appraisals being linked to revalidation. However, in the Consultant Career Grade questionnaire survey, some issues were raised. Some Consultants did not think that appraisals would effectively support the revalidation process (31.2%). For example, one stated that appraisal “depends ultimately on the honesty of the appraisee except in gross misconduct and aberration” (*Consultant Anaesthetics 040*). Another pointed out that “there will be too much

variation in the rigorousness of the appraiser for this to be a fair support to revalidation" (*Consultant Forensic Psychiatrist 417*).

Many respondents have doubts that appraisals will link to revalidation. One Consultant has pointed out that it is important to distinguish between appraisal and revalidation, stating that "appraisal is identifying needs of appraisee, and revalidation is allowing practitioner to continue practicing. These two have fundamentally different objectives" (*Consultant 039*). There is a similar concern that "revalidation is intended to identify doctors who are in some way under-performing or for some reason are a potential risk to the public. The appraisal process is incapable of doing this in its present form" (*Consultant 014*). One respondent claimed that there were inadequacies in the current appraisal process as it has "no evidence base, difficulty if investigated to prove in practice" (*Consultant 030*). Therefore, it is unlikely to achieve support for revalidation. Some suggested that the current appraisal process seems pervaded by politics and too greatly influenced by the government's agendas. One respondent argued that "the appraisal and revalidation process is political cosmetics" and will not bring any change within the NHS (*Consultant 104*).

### **Problems of appraisal implementation**

There are problems with the implementation of appraisal which were identified by the respondents in the questionnaire survey. These have been categorised as issues of fairness, efficiency, feedback and reliability, validity and relevance of the data. Evidence from respondents showed that there was a lack of equitable chances to access training, preparation time and appraisal across different clinical directorates. One respondent stated that in his/her Trust "staff grade appraisal has been very slow to start and has still not been completed. As usual Associate Specialists receive better treatment than Staff Grade, and Trust Grade" (*Associate Specialist 1293*). One respondent said, "I feel that the Trust is more enthusiastic when it comes to consultants and juniors, but the welfare of middle grades is grossly neglected" (*Staff Grade 0642*).

The evidence shows that the professionals feel it difficult and unfair to be appraised by their superiors; for example, relationship difficulties with the appraiser may affect the appraisal process. The unfairness at appraisal was described by one respondent, who felt their experience was "terrible, totally demoralized" because there was no listening, just "simply asked to obey them" (*Staff Grade Paediatrics 0863*). The respondents suggested that the lack of a standard format and structure might create unfairness between Consultant and Non-Consultant Grades. For example, "the Trust is using the same appraisal process set for consultants to appraise the NCCG doctors" (*Staff Grade 0937*).

There were problems of efficiency in the appraisal process identified by the respondents as inadequate support structures provided by management, lack of clear procedures, absence of a practical guide and inconsistencies in the system. Some respondents addressed lack of facilities support from the Trust, such as "lack of appropriate IT tools and support" (*Associate Specialists 0892*): "I am not sure for certain as to whether the process is likely to change circumstances within the Trust as limited resources and constraints within the infrastructures determine the ability to fulfil expectation for the future" (*Staff Physician 0799*).

Survey data have shown that there is a need to improve management education about appraisal. Some "senior staff are still unclear who should do appraisal" (*Staff Doctor 0883*) and "many clinical directors haven't got a clue about SAS jobs and commitments" (*Associate Specialist 0908*). Consistency in the conduct of appraisal was raised and there was evidence from the survey of a lack of formal training in conducting appraisals, for both appraisers and appraisees. One respondent recommended, "Standardise approach to training appraisers and appraisees. Ensure everyone is appraised by someone they can relate to and who understands their professional discipline" (*Staff Grade 0798*), and the Trust should be "considering equal opportunities" for training for all medical doctors in appraisal" (*Staff Grade 0817*).



### **Reliability, validity and relevance of data**

The reliability, validity and relevance of the data collected or needed and the difficulties of obtaining such data were commented on by both Grades. Concern was expressed about the lack of cooperation between different departments and lack of resources in accessing the necessary data. Non-Consultant Grade respondents showed considerable concern about the reliability and relevance of the data presented and discussed at appraisal. There were comments that there was no written evidence available in some Trusts. One stated, "No written evidence provided, but the rota of my job clearly demonstrated that I cannot gain certain CPD points or experience for furthering my professional knowledge" (*Staff Grade 0939*).

Some of the respondents expressed the opinion that insufficient evidence had been given in appraisal. The data contained inaccuracies and some of the data was missing, because of a lack of cooperation with other departments. One respondent claimed that it was "impossible to get accurate information from clerks, regarding numbers of patients seen etc, still very minimum IT within clinical scenario for audit" (*Associate Specialist 0942*). Another remarked: "The relevant information seems to have reached only certain parts of the medical staff" (*Staff Grade Psychiatrist 0972*).

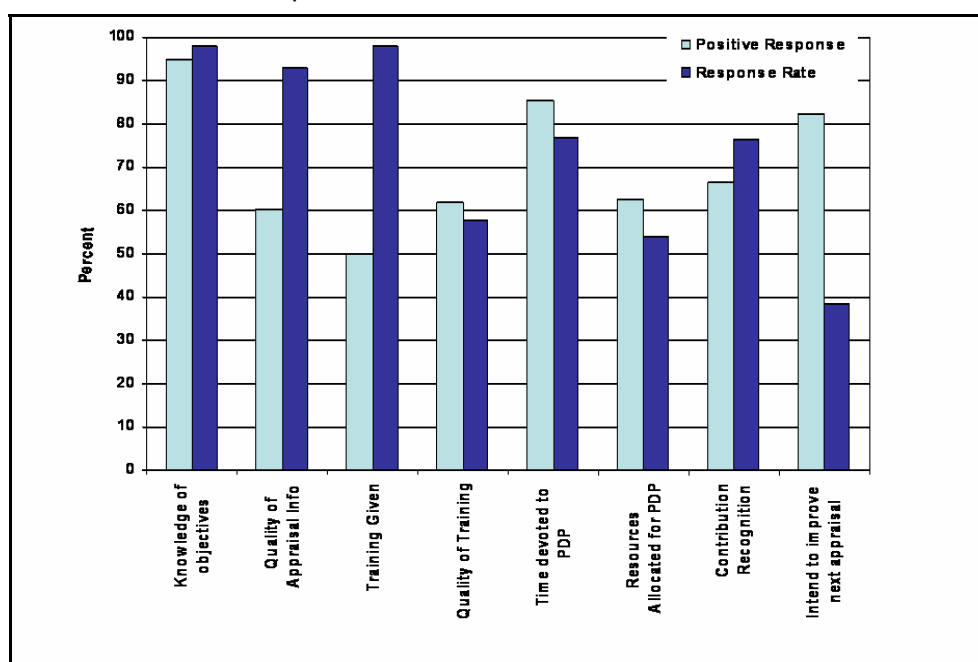
### **Strengths and weaknesses of the appraisal process**

There were comments from both career grades in the survey that appraisal was an imposition of a form of control from government, a 'paper exercise' and an 'expensive exercise,' without being tailored specifically for the circumstances of the medical professional. One respondent pointed out that current appraisal "was approached as a routine procedure. Not very person-specific, fairly superficial" (*Staff Grade A&E 1264*). Therefore, appraisal has been shown to be too weak as an effective control tool to monitor professionals' performance, because of its current inadequate design and lack of formal structure. One respondent described it as "lots of paper work, and centrally-planned but poorly coordinated" (*Staff Grade Paediatrician 1048*).

Another respondent questioned the effectiveness of appraisal and commented, "It felt like rubber stamping and felt an unnecessary use of time, given that we also have job planning, peer review" (*Associate Specialist 0606*). This evidence may suggest that if any real benefit is to accrue from the appraisal process then there is still much adjustment to be made. One respondent pointed out that the current appraisal process is "very much service oriented, not much place for career improvement" (*Staff Grade 1121*). Therefore, the current appraisal process might create stress and suspicion among the professionals.

### **Changes in professional medical behaviour in response to appraisal**

The answers to some key questions are summarised in Fig.1 in which both Grades have been aggregated with the relative response rates included.



**Figure 1.** Aggregated summary of factors potentially relevant to the change process

One impression gained from analysis of the questionnaires is that appraisal has produced little change except in the lowering of staff morale. The analysis does not produce significant evidence that appraisal has an impact on the improvement of professional behaviour. This might indicate that the way the current appraisal process is conducted lacks incentives for the professional to change.

One respondent claimed, "I get appraisal every year. Every year I was promised on associate specialist grade (for the last three years) still nothing happened. It is the same talk, the same time waster" (*Staff Grade Cardiologist 1421*). Some respondents expressed less enthusiasm for future appraisals, with one claiming that "I wouldn't put much effort in as it [appraisal] as just a paper exercise, a formality really. It is [a] total waste of time" (*Staff Grade Psychiatrist 1386*), and another "wouldn't bother preparing anything, as all my evidence was glossed over and not heard" (*Associate Specialist 1150*). On the other hand, some respondents said they would spend more time and effort to prepare for it, but doing so by spending less time on patients. As one Consultant claimed, "I would cancel several clinics rather than waste my own time on this" (*Consultant Dermatologist 306*).

## 4.2 Interviews

The empirical results from interviews revealed six major themes: professional development; recognition of poor performance; the link to revalidation; appraisal as a tool of change; changes in organisational and medical practices in response to appraisal; and patient safety.

### **Professional development**

One function of appraisal is to "set out personal and professional development needs" and "agree plans for them to be met" according to the *Department of Health* (1999: 1), which fails to mention the allocation of resources to implement these plans, whilst mentioning resources for service needs. In Non-Consultant interviews, there were many interviewees who considered that appraisal in itself was good and it had provided opportunities for professional development, though some of them admitted that they did not have much experience of it, even though appraisal has been good practice in commerce and industry for a long time. One pointed out that appraisal was "a positive experience. I learned from it", but further suggested that there was a bureaucratic system which appeared to obstruct full professional development in the Trust. This interviewee went on to say, "when I did want to develop I make a report of that. That goes through to the Medical Director, who then reports to the Chief Executive, so there is record of where you might like to develop. Whether your Trust will support you in doing that provide you with study leave, provide you with finance are very questionable. And there is a threat out there that they will be less inclined to do that unless they see specific service benefits for the Trust" (*Associate Specialist Anaesthetist 001*).

However, there were suggestions that appraisal is not necessary to enable professional development to happen: "in the current climate, training needs of NCCGs are not a priority, not in our Trust. So resource issues mean that even if training needs are identified, there isn't any help in developing further to meet them. It's all still driven by the individual. There is no particular allocated time; we just fit things in between our clinical commitments" (*Associate Specialist 006*); "From the appraiser's point of view, appraisal was just about conducting the appraisal - a tick in the box. Afterwards the onus is all on the appraisee to take any action" (*Associate Specialist Paediatric Cardiology 008*).

Similarly, the majority of Consultants interviewed believed that appraisal could help individual professional development. There were, however, some interviewees who expressed negative views of the appraisal process, which does not encourage a professional to appreciate that appraisal could be a useful tool to improve professional development. One has said that appraisal "is just a ritual, but I wonder whether I should take steps to make it a useful process or or or if I should shrug my shoulders and let whoever it is go through the rigmarole of it." (*Eye Cancer Consultant 002*). Another Consultant questioned the cost/benefit of appraisal for development because "most people are professional enough to do their own stuff to keep up to date and it is a long process - the paper work, and the appraisal is time consuming"; it is "a sledgehammer to crack a nut" (*Consultant 003*).

### **Recognition of poor performance**

Evidence from Non-Consultant Grade interviews shows that, although many interviewees considered that appraisal is a good thing for monitoring professional performance, there were strong suspicions that there were political forces driving appraisal. The medical adverse events which were exposed in the Shipman Inquiry (2002) and the Bristol Inquiry (2001), put the government under pressure to increase scrutiny from the outside, because individual doctors had abused traditional self-regulation. However, from an interviewee point of view many felt that external regulation has been imposed because of isolated cases. It was implied that appraisal might not detect incompetent doctors, but it could demoralise good ones: "One GP like Shipman killed some of his patients, so then all the GPs are subject to control by the government" (*Associate Specialist 004*).

The current appraisal process has led to some cynicism over its real purpose, as one critic suggested, "the process is a meaningless paper exercise to allow the government to claim that it is 'doing something' to prevent another Shipman/Ledward etc." (*Consultant 14*). Another was "not very sure appraisal would do anything to catch the rogues" (*Anaesthetics 009*).

Several interviewees questioned whether the NHS would benefit from using appraisal to detect poor performance. One interviewee voiced concerns over this function of appraisal: "I have serious doubts. I think appraisal plays a limited role. Good practitioners will continue to practice well, and it would not make much difference to those who don't" (*Associate Specialist 008*). One criticism offered was that "the whole process still relies on individual's honesty. Appraisal itself would not pick up poor performance, like Shipman" (*Rheumatology 007*). The interview respondents in the Consultant Grades appeared to show similar concerns to Non-Consultant Grades in that they felt appraisal would not be the most efficient way to detect poor performance.

### **Link to revalidation**

Whilst both Grades expressed some positive views about appraisal in general, there was more caution about revalidation and its link to appraisal conduct. One expressed concern that there was a lack of clear differentiation between Consultants and Non-Consultants in the documentation. One interviewee pointed out that appraisal and validation were not very discriminatory in revealing professional performance and that "to me they just a paper exercise...we have got... bits of paper to say I have been appraised, I have been revalidated. But nobody really knows what I am doing or how good I am. I could just be good at filling in papers. You could be a really bad doctor doing all sorts of things, but it wouldn't show in the appraisal" (*Associate Specialist 003*). The validity of linking appraisal directly to revalidation has been questioned.

Several interviewees viewed the notion of revalidation as problematic in practice in the NHS, so there was anxiety and fear about the appraisal process being linked to revalidation. Although interviewees were aware of the nature of appraisal practice in industry as a kind of normative process, in the medical setting appraisal performance is linked to revalidation and therefore the licence to practice. One interviewee admitted, "I suppose that most professionals feel it's a bit threatening to some extent. But most of the department fear naturally." (*Associate Specialist 004*).

### **Appraisal as a tool of change affecting job satisfaction, motivation and commitment**

Interviewees from both Grades commented on the many interventions from government in recent years, for example, target setting, the National Institute for Clinical Excellence, Commission for Health Improvement and the appraisal process, intended to effect changes in the health care system. These have had considerable negative impacts on doctors' job satisfaction, motivation and commitment.

One interviewee provided an example about the new consultant contract and incentives, which is a very difficult area:

"There was a contract negotiated, the doctors rejected it, the government decided that what they would do is have a system of either you chose the

contract or you could have incentive payments and different styles of incentive payments in different places" and "we have got three or four different schemes now, to implement in a climate where everybody's pretty fed up about what's happened in the last few months and over the pay scheme which was in a climate of a lot of anxiety amongst medical staff about the attitude of the government to them, the behaviours managers were exhibiting around things like targets and that sort of thing, so they knew the government had instructed them, but they still didn't feel that managers were doing what they wanted them to do, and then there is this whole piece of work around morale going on, around the whole compact between the public and medicine."

*(Director of 'New Ways of Working')*

This shows the frustration and dissatisfaction with the way in which government initiatives affect management processes and ignore the needs of the medical profession. This has led to a feeling of disempowerment and unilaterally imposed external scrutiny. Therefore there is an impact on morale, reducing commitment and damaging motivation. One has claimed that the appraisal process had a personal negative impact because "I do an awful lot, yet it is never enough for people. They always want to get as much clinical work out of you as possible. And you never get any praise for anything" (*Associate Specialist 003*). One respondent argued that those government interventions have to some extent changed the current medical culture for the worse, for example "The junior doctors have much less commitment to the patients. It is more of a job to them." (*Rheumatology 007*).

### **Changes in organisational and medical practices in response to appraisal**

Most Non-Consultant and Consultant Career Grades do not believe that the appraisal process has led to much significant change in the organisations, but that there was some impact on professional performance. However, the data failed to provide clear evidence to illustrate this. Some findings indicate that there was a negative response to appraisal, as its supposed purpose of performance control has challenged the medical professionals' traditional self-regulation. Most of the interviewees of both grades have recognised that current trends in the NHS control system, representing shifts in autonomy in terms of regulation, revalidation and other forms of surveillance are detrimental to the notion of professionalism, because self-regulation is one of the essential characteristics of professionalism and professional bodies.

For example, although some interviewees express no objection to public accountability, "if it is done in the right spirit" (*Rheumatology 007*), some consider that "it has all been taken too far" (*Paediatrics Specialist 010*). One said, "they seem to have gone from one extreme to the other. Self-regulation has its faults but the current swing to openness and public accountability has not been properly thought through" (*Rheumatology 007*). In expressing apprehension over increased scrutiny of the medical profession, one specialist was "very concerned about that responsibility being taken from Colleges" (*Consultant 003*).

Some respondents suggested that replacement of traditional self-regulation by peer supervision, which might be called a low trust form of regulation, with a tight control form of surveillance may cause damage to morale. One respondent suggested that, "It's all about 'Watch out, for you might doing wrong!' but in return you never get praise for what you do good [sic]. You probably never did anyway. Mistakes are more likely to happen because you are tired and you are rushed. It is wearing away one's very morale" (*Associate Specialist Dermatologist 003*). Currently accountability is now being taken outside that realm and is being introduced in a more formal way, challenging traditional forms of professional assessment and self-regulation, by doctors keeping an eye on and between each other. The interviewee expressed dissatisfaction with the regulation of doctors being opened up to the public, although appraisal was considered to be "a part of the regulation. It is sort of self-regulation because [it is] within the medical profession itself. It is not somebody come from outside" (*Associate Specialist 002*). One respondent believed that the increased scrutiny from outside could be detrimental to the patient. Whilst many recognised the problems of self-regulation, there was widespread resistance to the present approach.

### **Patient safety**

Interview evidence from both grades has suggested that, in theory, appraisal is a good thing, if it is applied properly. It could be a positive influence for promoting certain changes within the organisations, for example, it could improve patient care by enhancing professional

reflection (Myerson, 2000). However, there was evidence that the current appraisal process conducted in some hospitals might be less effective in facilitating organisational changes. One interviewee clearly stated that the mechanism of appraisal applied in the organisation created conflict between doctors and management, and therefore any changes would be unlikely to take place within the organisation. The respondent commented, "'Performance by results' is ridiculous. Will reward the large Trusts with more staff and not those trying to improve. By increasing emphasis on targets rather than patients. Clinicians have less of [a] voice than previously, 'Us and them' is increasing between clinicians and managers" (*Anaesthetics Specialist 009*).

One interviewee believed that "appraisal is being used in the NHS in the wrong way. It was intended to improve the clinician's performance through a positive supportive developmental process. Instead it is being used to target individuals with negative criticism with no positive impact on patient care - may even be negative" (*Paediatrics Specialist 010*). Another interviewee believed that, "if the doctor's further development needs are identified and if these are then met it will improve their clinical practice and expertise, and improve the quality of clinical care provided for patients" (*Ophthalmology Specialist 005*). One interviewee claimed that appraisal can help to improve patient care "theoretically" and, "if the appraisal process is done properly, it can impact patient care, but most importantly it can impact on doctors themselves" (*Consultant 001*).

## 5. Discussion

The questionnaire survey and semi-structured interviews were employed to explore the appraisal process and its potential impacts, and provided data to address the key questions of appraisal implementation: how far doctors believe that appraisal linked to revalidation will strengthen the control of their performance to identify problem doctors and to improve patient safety, and how far doctors believe that it is possible to closely regulate their profession, given the nature of the work they perform.

The evidence from the questionnaires, in particular, seems to indicate that generally the quality and effectiveness of the implementation is less than satisfactory. Respondents indicated that the current appraisal process has little significant influence on enhancing their job performance, and similar findings can be found in Herbert *et al.*'s studies as early as 1965, in determining the effect of performance appraisal and job performance. The data also indicated that transplanting the concept of appraisal from the private sector to the NHS has been done with little regard to the resources required and its unintended consequences.

The empirical findings from questionnaires and interviews indicate that there is no clear link seen between appraisal and revalidation. Similar findings can be seen in Cornish's study *GP Appraisal in Hertfordshire* (2006: 102), which concluded "there is still uncertainty surrounding the role of appraisal and revalidation," and Adams *et al.*'s (2004) study of the appraisal process for general practitioners. There is no conclusive evidence about how the current appraisal process and its link to revalidation are suitable for improving professional accountability and controlling and monitoring performance. It is clear from the evidence that evaluation of medical performance for the purpose of continuing professional development was felt by many responders to be at odds with evaluation as a judgment of competence, which was seen by many as potentially threatening to their livelihood. The evidence has shown that "an increasingly complex system for ensuring accountability can undermine the professionalism it is supposed to safeguard. And doctors may feel less inclined to behaviour altruistically if they are excessively scrutinised" (Rosen & Dewar, 2004: 46).

The design and implementation of the current appraisal process has created anxiety and resistance, with informants calling it 'a paper exercise', 'box-ticking games', 'ritualistic', 'mock ritual' and a bureaucratic process. Evidence in this study showed some doctors believed that the extra paperwork would lead to the practice of more 'defensive medicine' and 'reward bad practice', (MORI, 2005) by encouraging them to spend less time with patients in order to fill in the paperwork that would secure their future (Levenson, *et al.*, 2008). Some of the findings here resonate with McGiven & Ferlie's (2007: 1362) study on consultant appraisal, with the authors regarding the process as 'box-ticking games.'

The empirical findings in this study are equivocal as to whether the exclusive use of appraisal will be sufficiently robust for the anticipated improvements in NHS control systems. It is suggested that the appraisal process is a tangible example of the intangible disciplinary power

of surveillance. Such control mechanisms reflects an ambivalent structure of governance involved in the process by means of 'soft bureaucracies' (Courpasson, 2000).

## 6. Conclusion

This paper began with the need for improved control and the enforcement of self-regulation of doctors in the UK NHS, aiming to investigate doctors' responses to the recently introduced appraisal process, and to identify more appropriate controls of medical performance. One mechanism that the UK government believes, both implicitly and explicitly, may be instrumental in making these improvements to medical performance and control of doctors, is the imposition of mandatory individual appraisals leading to revalidation. This paper has evaluated the efficacy of this appraisal process within the context of seven major government policies since 1983, which were influential in shaping NHS control mechanisms.

The main conclusions are, firstly, that appraisal may generate the disciplinary power of surveillance to influence doctors to change behaviours with the intention of controlling their performance, and making it accountable and predictable. Secondly, the appraisal process was seen by many respondents as a means of the State wresting from doctors the powers of traditional medical autonomy, freedom and self-regulation. The State has never been able to control doctors since the NHS was founded. Now it seeks to replace professional autonomy with responsible autonomy. Thirdly, appraisal linked to revalidation as a policy prescription has been oversold in the current climate of increasing public demands on the NHS and reduced confidence in doctors. Fourthly, the linking of appraisal to revalidation has begun to create confusion, fear, uncertainty about its future implementation, reduction in morale and increased resistance to change. Therefore, the fear of sanctions - in the form of reluctance to re-licence in cases of failure in appraisal - is likely to result in outcomes that demoralise and generate resistance. Hence, the methodological boundaries of appraisal and revalidation should be made contiguous.

The findings contribute empirical evidence for developing a more theoretically based understanding of the current appraisal process and its defects in regulating and controlling doctors' performance to make professional accountability more responsive, with the ultimate aim of improving health service delivery, especially patient safety. However, the main limitations of this study are that the effectiveness of decreasing self-regulation and increasing bureaucratic control through appraisals cannot be convincingly evaluated here, because of the sample size and because of the absence of any ethnographic observation. Yet indications from the empirical work suggest that many respondents are anxious, suspicious, and lack confidence in the appraisal process and that some of them may be willing to circumvent it. Further research is merited using larger samples and controlling for poor appraisal implementation.

## References

- Adams A. Illing J. Zwanenberg T. & Walker C. (2004) Will GP appraisal undermine practice professional development plans? *Clinician in Management*. 12: 187-90.
- Buckley, P. (2007) 'Professional regulation: the implications of 'trust, assurance and safety - the regulation of health professionals in the 21st century,' *Clinical Risk*, 13: 99-102.
- Coburn, D. Rappolt, S. & Bourgeault, I. (1997). Decline vs. retention of medical power through restratification: an examination of the Ontario case. *Sociology of Health & Illness* 19 (1): 1-22.
- Cornish, M. (2006) GP Appraisal in Hertfordshire: Quality Assurance of the GP Appraisal Scheme in Hertfordshire, Bedfordshire & Hertfordshire, Local Medical Committee.
- Courpasson, D. (2000) Managerial strategies of domination: power in soft bureaucracies, *Organization Studies*, (22): 141-61.
- Dent, M. (2005) Post-New Public Management in public sector hospitals? The UK, Germany and Italy. *Policy & Politics*, 33 (4): 623-636.
- Donaldson, L. J & Gray, J. A. (1998) 'Clinical governance: a quality duty for health Organizations,' *Quality in Health Care*, 7 (Suppl): S37-S44.
- Farbey, B. Land, F & Targett, D. (1999) The moving staircase - Problems of appraisal and evaluation in a turbulent environment. *Information Technology & people*, 12 (3): 238-252.

- Ferlie E. (1998). The new public management in the United Kingdom: origins, implementation and prospects. Paper for the 'managerial reform of State' International Seminar, organized by the Ministry of Administration and State Reform, Brasilia, Nov, 1998.
- Fierlbeck, K. (1996), Policy and ideology: the politics of post-reform health policy in the United Kingdom. *Int. J. Health Serv.* 26(3): 529-546.
- Flynn, R. (1992), Structures of Control in Health Management, London: Routledge.
- Greener, I. (2001) "The ghost of health services past" revisited: comparing British health policy of the 1950s with the 1980s and 1990s. *International Journal of Health Services*, 31 (3): 635-646.
- Greener, I. (2004) Health service organisation in the UK: A political economy approach, *Public Administration*, 82 (3): 657-676.
- Greener, I. (2008). Decision Making in a Time of Significant Reform: Managing in the NHS, *Administration & Society* 2008 40: 194-210
- Hann. A. (2000) Analysing Health Policy. Ashgate.
- Harrison, S. (1991). Working the markets: purchased/provider separation in English health care, *International Journal of Health Services*, 21(4): 625-35.
- Herbert, H. Meyer, E. K. & French, Jr. J. R. P. (1965) 'Split roles in performance appraisal' *Harvard Business Review*, (43): 123-9.
- Hunter, D. (1995) Recent developments in health policy: an overview. *The Asian Journal of Public Administration*, (17): 22-42.
- Klein, R. (1995, 3rd ed). The New Politics of the NHS. Longman, Harlow.
- Klein, R. (1998) 'Why Britain is reorganising its National Health Service – yet again'. *Health Affairs*, 1998
- Klein, R. (2000). *The New Policy of the NHS*, Harlow: Longman.
- Kuhlmann, E & Allsop, J. (2008) Professional self-regulation in a changing architecture of governance: comparing health policy in the UK and Germany. *Policy & Politics*, 36 (2): 173-89.
- Levenson, R. Dewar. S. & Shepherd. S. (2008) *Understanding Doctors: Harnessing Professionalism*. King's Fund.
- Lewsi, J. (1998). The Medical profession and the State: GPs and the GP Contract in the 1960s and the 1990s, *Social Policy and Administration*. (32): 132-60.
- McGivern, G and Ferlie, E. (2007). Playing tick-box games: Interrelating defences in professional appraisal. *Human Relations*, 60 (9): 1361-1385.
- MORI (2005) Attitudes to medical regulation and revalidation of doctors: Research among doctors and the general public. Research study conducted for Department of Health.
- Nettleton, S. Burrows, R & Wat, I. (2008) Regulating medical bodies? The consequences of the 'modernisation' of the NHS and the disembodiment of clinical knowledge, *Sociology of Health & Illness* 30 (3): 333-348.
- North N., Lupton C. & Khan P. (1999) Going with the grain? General practitioners and the new NHS. *Health and Social Care in the Community* 7 (6): 408-416
- Ong, Bie Nio. Boaden, M. & Cropper, S. (1996), Analysing the medicine-management interface in acute trusts. *Journal of Management in Medicine*. 11 (2): 88-95..
- Powell, M. (1999). The Third Way in the Health Care Sector: New Labour and Third Way in the British National Health Service, *International Journal of Health Services*, 29 (2): 353-370.
- Strong, P. & Robinson, J. (1990). The NHS Under New Management, Open University Press, Milton Keynes, Philadelphia.
- West, P. (1998). Market—what market? A review of Health Authority purchasing in the NHS internal market. *Health Policy*, 44, 167-183.
- Woods, K. J. (2002). A critical appraisal of accountability structures in integrated health care systems. Paper prepared for the Scottish executive Health Department Review of Management and Decision making in NHS Scotland.
- Van Zwanenberg T. (2004). Revalidation: the purpose needs to be clear. *BMJ* 2004; 328: 684-686

#### Government Documents

DH, *Statistical Bulletin*, 1994

Department of Health (2007). Trust, Assurance and Safety – The Regulation of Health Professionals in the 21st Century, HM Stationery Office, London.

Matthews QC (2004). Committee of inquiry to investigate how the NHS handled allegations about the performance and conduct of Richard Neale, Department of Health. London.

Ritchie, J. (2000). Inquiry into Quality and Practice within the National Health Service arising from the Actions of Rodney Ledward, HM Stationery Office, London.

Smith, J. (2005). The Shipman Inquiry: Fifth Report; Safeguarding Patients: Lessons from the Past – Proposals for the Future, HM Stationery Office, London.

Department of Health (Kennedy, 2001). Bristol Royal Infirmary Inquiry. Learning from Bristol: the report of the public inquiry into children's heart surgery at the Bristol Royal Infirmary 1984-1995. London: Department of Health.

Nigel Pleming, QC (2005). The Kerr/Haslam Inquiry: Full report, London: Department of Health.

Redfern, M. et al, (2000). The Report of The Royal Liverpool Children's Inquiry. HM Stationery Office, London.

Department of Health. (1989). Working for Patients. London: HMSO.

DHSS (1983), The NHS Management Inquiry (The Griffiths Report), HMSO, London.

Secretary of State for Health. (1998). A first class service: quality in the new NHS. London: Department of Health.

Donaldson, *et al*, (2000). An organisation with a memory report of an expert group on learning from adverse events in the NHS. London: Department of Health.

Department of Health. (2000). The NHS plan: a plan for investment, a plan for reform. London.

Department of Health. (1999). Supporting doctors, protecting patients: a consultation paper on preventing, recognising and dealing with poor clinical performance of doctors in the NHS in England. London.

General Medical Council (1995). Duties of a doctor: good medical practice. London: GMC.



# Current Usability Metrics for Mobile Computing Evaluation

**Azham Hussain and Maria Kutar**, *Informatics Research Institute, University of Salford*

## Abstract

Usability metrics are important elements for measuring whether an application is usable or unusable. Literature on how to measure usability is limited in the area of mobile computing. In this paper, we review, categorise, and discuss usability metrics obtained from 26 studies published in core human-computer interaction (HCI) journals. The analysis produces usability metrics for evaluating mobile computing (as distinguished from usability metric used in desktop computing). We also explain the methods used by researchers to create usability metrics and describe several issues and challenges in developing usability metrics in mobile computing.

## 1. Introduction

Usability metrics are essential elements in software measurement; they ensure that the application is accurate, increase speed, and ensure the safety of the user from strain injury (Ahmed, Mohammad, Rex, & Harkirat, 2006). Mobile device applications such as news alert, weather forecasting, and entertainment have become increasingly popular and well accepted. The fast growth and high demand for mobile applications has encouraged researchers to undertake studies on potential developments in this field.

Many advantages can be obtained when conducting usability testing, such as driving down production costs, improving sales, enhancing brand loyalty, and providing access for all in the public sector (Abrahams, 2008). However, conducting usability testing will be rather difficult without usability metric, which provides guidelines for the usability evaluation process. Hence, it will be beneficial to have usability metric either for desktop computing or mobile computing before conducting usability testing.

Mobile computing can be described as software systems operating on mobile devices (Zhang & Adipat, 2005). This classification is divided into two types: horizontal applications and vertical applications (Thomas & Mark, 2003). Horizontal applications are general and adaptable to a wide range of users and applications, such as music, e-mail, web browsers, and file transfer. On the other hand, vertical applications are aimed at a specific type of user or application, such as financial, marketing and advertising, education, and emergency applications. This wide range of mobile applications creates many research opportunities for researchers.

Clear guidance on how to measure usability is limited (Hornbæk & Law, 2007), and much more so in the area of mobile computing itself. The novelty of mobile computing and the unique features of mobile devices are the main challenges in usability measurement activity. Recent advances in technology like global positioning system (GPS) receivers embedded into mobile phones create new challenges in the human-computer interaction (HCI) area. Many traditional usability metrics have been created intentionally for desktop applications; however, these metrics may not be directly applicable to mobile applications (Zhang & Adipat, 2005).

This paper aims to review the literatures on usability evaluation methods and current practice on modelling usability measures. We also identify the current usability metrics for mobile computing and explain the challenges in measuring usability. The next section provides a review of usability metrics, followed by an outline of the research strategy employed in this study. Finally, we explain and discuss the results obtained from this research and draw conclusions.

## 2. Related studies

Mobile computing imposes different usability challenges compared with still technology—for example, in terms of designing usable interfaces (Ballard, 2005); how usability is measured (Kjeldskov & Stage, 2004); and the relationships between the technology, work tasks, and the context of work (Steinar & Fredrik, 1999). Understanding the usability of mobile computing has been widely discussed—for example, in the navigation of complex information on small screens (Björk, Johan, Peter, & Lars Erik, 2000); tactile feedback (Brewster, Faraz, & Lorna, 2007); techniques for assessing mobile usability (Kjeldskov & Stage, 2004); Barnard et al., 2005; and (Minna et al., 2007); texting (MacKenzie & Soukoreff, 2003); and how to conceptualise mobile usability (Dong-Han et al., 2006). Usability evaluation methods refer to the techniques employed to carry out usability evaluation, such as usability testing, focus groups, and interviews. All of these methods have been used by many researchers to evaluate usability, and each method has advantages and disadvantages, depending on the objective of the study. Different evaluation methods have emerged and contributed to the evolution of usability evaluation, giving software development organisations a wide collection of techniques that fit specific development projects (Scapin & Law, 2007).

A number of models for usability measurement are available for reference—for instance, Quality in Use Integrated Measurement (QUIM), developed by Ahmed et al. (2006). QUIM is a consolidated model for usability measurements and metrics; it is also appropriate for users who have no or little knowledge of usability. The model consists of 10 factors that are subdivided into 26 criteria. For the measurement of these criteria, the model provides 127 metrics. The model is used to measure the actual use of working software and identify the problem; however, it is not optimal yet and needs to be validated.

On the other hand, Metrics for Usability Standards in Computing (MUSiC), developed by Bevan and MacLeod (1994), is another project concerned with defining measures of software usability. MUSiC is integrated into the original ISO 9241 standards, which are effectiveness, efficiency, and satisfaction. Examples of specific usability metrics in the MUSiC framework include user performance measures, such as task effectiveness, temporal efficiency, and length or proportion of productive period. However, a strictly performance-based view of usability cannot reflect other aspects of usability, such as user satisfaction or learnability. Software Usability Measurement Inventory (SUMI), developed by Kirakowski and Corbett (1993), is a part of the MUSiC project. SUMI was developed to provide measures of global satisfaction for five additional usability areas: effectiveness, efficiency, helpfulness, control, and learnability. Another MUSiC project related to software tool development is Diagnostic Recorder for Usability Measurement (DRUM) (Drummond & Themessl-Huber), developed by Macleod and Rengger (1993). This project focuses on the analysis of user-based evaluations and delivery of these data to the appropriate party, such as a usability engineer. The Log Processor component of DRUM is the tool concerned with metrics. It calculates several different performance-based usability metrics, including task time; snag, help, and search times; effectiveness; efficiency; relative efficiency; and productive period.

Subsequently, another model dealing with the analysis of the quality of use for interactive devices was introduced by Macleod & Rengger (1993)—namely, the Skill Acquisition Network (SANE). This approach assumes a user interaction model that defines user tasks, the dynamics of the device, and procedures for executing user tasks. Specifically, a task model and a device model are simultaneously developed and subsequently linked. Afterwards, user procedures are simulated within the linked task-device model. A total of 60 different metrics are described in this framework, 24 of which are concerned with the quality measures. Scores from the latter are then combined to form a total of five composite quality measures, including efficiency, learning, adaptiveness, cognitive workload, complexity, and effort for error correction.

The International Organization for Standardization (ISO) is an international standard-setting body composed of representatives from various national standards organisations. The ISO has developed over 17,000 international standards for a variety of subjects, and 1,100 new ISO standards are published every year (ISO website, [www.iso.org/iso/home.htm](http://www.iso.org/iso/home.htm)). Most literature in HCI employs ISO 9241-11 for usability measurement (Hornbæk & Law, 2007). Table 1 lists the ISO standards related to HCI. ISO 9241-11 specifically addresses the definition of usability measurement and has been thus chosen as the foundation for the framework in this study. In another study of usability measurement dimensions, Constantinou and Dan (2007) found the

most common characteristics in usability evaluation are effectiveness (62%), efficiency (33%), and satisfaction (20%). These three characteristics reflect the ISO 9241-11 standard, as they measure attributes for that standard.

**Table 1.** ISO Standard related to measurement.

<b>Usability in ISO Standard</b>	<b>Description</b>
ISO 9241-11 (1998)	Identifies efficiency, effectiveness, and satisfaction as major attributes of usability.
ISO/IEC 9126-1 (2001)	Defines the standard as software quality attributes that can be subdivided into five different factors, including understandability, learnability, operability, attractiveness, and usability compliance.
ISO/IEC 9126-4 (2001)	Defines the related concept of <i>quality in use</i> as a kind of higher-order software quality attribute.
ISO/IEC 14598-1 (1999)	A model for measuring quality in use from the perspective of internal software quality attributes

### 3. Research Approach

This section explains the steps taken to review the literature. A systematic literature review (SLR) method will be used in this phase to ensure that previous studies on usability evaluation relevant to this study are considered. The journals selected in this study were the top journals in HCI from 2006 until 2008. The total of 409 journal papers was reviewed based on keywords "usability", "evaluation" and "metric". Only 26 out of 409 journal papers were selected for further review in obtaining the guidelines and usability metric for mobile application development. Table 2 below describes the journal papers that were reviewed.

**Table 2.** Journal papers reviewed.

<b>Journal</b>	<b>Year</b>	<b>Candidate</b>	<b>Selected</b>
<i>ACM Transactions on Computer-Human Interaction</i>	2006–2008	54	8
<i>Human-Computer Interaction</i>	2006–2008	36	2
<i>International Journal of Human-Computer Interaction</i>	2006–2008	97	5
<i>International Journal of Human-Computer Studies</i>	2006–2008	222	11
Total		409	26

The review is based on a conception of usability that is similar to ISO 9241-11 (1998) and Bevan and MacLeod (1994). This conception merely discusses studies related to usability evaluation rather than the broad concept of usability. We analyse the quality characteristics of each measure to ensure that no duplication exists. Interestingly, we found that most of the studies employed *effectiveness*, *efficiency*, and *satisfaction* as quality characteristics, which appear in ISO 9241 as well. Thus, we decided to make these three characteristics base guidelines and to designate others as sub-guidelines. Table 3 describes the most popular usability guidelines obtained from the literature. In the following section, we review the approaches that previous studies utilised to identify the issues that arise in creating usability metrics and discuss the results.

**Table 3.** Usability guidelines from the literature.

No	Guidelines	Description
1	Completeness	The extent or completeness of users' solutions to tasks.
2	Accurate	The accuracy with which users complete tasks.
3	Fewer errors or no errors	Errors made by the user during the process of completing a task.
4	Ease of data input	The data input process should be simple.
5	Ease of output use	The output should be very simple and accurate.
6	Ease to install	Application installation should be user friendly.
7	Response time	The system must respond in an appropriate time.
8	Simple	The application should be straightforward.
9	Time	The duration of tasks or parts of tasks.
10	Ease to learn	The user interface must be designed for the user to learn easily.
11	Application size	The space used by the application should be appropriate.
12	Battery power used	The battery power use by the application.
13	Wireless connectivity	The application should easily connect to a network.
14	Features available	Appropriate features available on an application.
15	Satisfaction with the interface	Measures satisfaction as the interface that users prefer using.
16	Provide support/help	The help information given by the application should be useful.
17	Safety	User should be safe and secure while using the application.

#### 4. Results and discussion

With reference to Table 3, we simplify the usability metric by removing the guidelines that are not relevant to interaction. Duplicate guidelines are omitted as well, to assure that they have no effect on other guidelines. Since this study focuses on mobile computing, we create several new guidelines related to mobility—for instance, "touch screen facilities," "safety while driving," and "automatic update"—which do not appear in existing literature. In Table 4, we present the results by categorising quality characteristics and their corresponding usability guidelines and metrics.

**Table 4.** Usability metrics for mobile computing.

Quality Characteristic	Usability Guidelines	Usability Metrics
Effectiveness	Simplicity	<ul style="list-style-type: none"> <li>• Ease of data input</li> <li>• Ease of output use</li> <li>• Ease of installation</li> <li>• Ease of learning</li> </ul>
	Accuracy	<ul style="list-style-type: none"> <li>• Accuracy</li> <li>• Absence of errors</li> <li>• Success</li> </ul>
Efficiency	Time taken	<ul style="list-style-type: none"> <li>• To respond</li> <li>• To complete a task</li> </ul>
	Features	<ul style="list-style-type: none"> <li>• Support/help</li> <li>• Touch-screen facilities</li> <li>• Voice guidance</li> <li>• System resource information</li> <li>• Automatic updating</li> </ul>

Satisfaction	Safety	<ul style="list-style-type: none"> <li>• While using the application</li> <li>• While driving</li> </ul>
	Attractiveness	<ul style="list-style-type: none"> <li>• User interface</li> </ul>

In the current literature, several issues arise with respect to creating usability metrics. For example, few studies employ experts to assess quality. Expert assessments are needed to evaluate subjective data and usability defects (Bevan & MacLeod, 1994). Hence, we would suggest employing experts to create usability metrics, as expert judgement is one of the essential methods in usability inspection. We also note that measures on how to use interfaces are rarely employed. Current measurements for interfaces focus more on the use of colours; layout and information structuring; consistency of the terminology; consistency of the interaction mechanisms; and how to fit in small-screen devices. Metrics for interface use will be different for mobile applications due to the novelty and size of mobile devices.

The measure of satisfaction is another important element in measuring usability, and questionnaires are one of the techniques employed to measure user satisfaction. However, many studies are not using validated questionnaires to assess satisfaction. A large number of questionnaires have been developed to assess the user's subjective satisfaction of the system and related issues. These include the Questionnaire for User Interface Satisfaction (QUIS), developed by Chin et al. (1988) and SUMI, developed by Kirakowski & Corbett (1993). Finally, we found that many studies combine both objective and subjective measures—for example, "learnability of an interface" and "time needed to master an interface into a single metric." We believe that such measurements should be separated rather than included in a single metric.

## 5. Conclusion

We have reviewed usability metrics employed in 26 studies published in core HCI journals, drawing attention to several issues relating to the production of these measurements. The metrics created in this paper could serve as alternatives to current usability metrics in mobile computing. In addition to the issues raised in this study, we intend, in future research, to explore other challenges in creating usability metrics, particularly in the area of interaction between humans and mobile applications.

## References

- Abrahams, P. (2008) 'The future of usability testing,' Bloor Research; available at: <http://www.it-director.com/business/compliance/content.php?cid=10495>.
- Ahmed, S., Mohammad, D., Rex, B.K., and Harkirat, K.P. (2006) 'Usability measurement and metrics: a consolidated model,' *Software Quality Control* 14 (2), 159–178.
- Ballard, B. (2005) *User Interface Design Guidelines for J2me Midp 2.0*. Lawrence, KS: Little Springs Design, Inc.
- Barnard, L., Yi, J.S., Jacko, J.A., and Sears, A. (2005) 'An empirical comparison of use-in-motion evaluation scenarios for mobile computing devices,' *International Journal of Human-Computer Studies* 62 (4), pp. 487–520.
- Bevan, N. and MacLeod, M. (1994) 'Usability measurement in context,' *Behaviour and Information Technology* 13 (1), 132–145.
- Björk, S., Redström, J., Ljungstrand, P., and Holmquist, L.E. (2000) 'POWERVIEW: using information links and information views to navigate and visualize information on small displays,' *Proceedings of the 2<sup>nd</sup> International Symposium on Handheld and Ubiquitous Computing*, Bristol, UK.
- Brewster, S., Chohan, F., and Brown, L. (2007) 'Tactile feedback for mobile interactions,' *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Washington, D.C., pp 159–162.
- Chin, J.P., Diehl, V.A., and Norman, K.L. (1988) 'Development of an instrument measuring user satisfaction of the human-computer interface,' *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Washington, D.C., pp. 213–218.

- Hornbæk, K., & Law, E. L. C. (2007) 'Meta-analysis of correlations among usability measures,' *Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems*.
- Dong-Han, H., Jeongyun, H., Fossick, P., Wong, W., Park, S., Chiwon, S., and Bradley, M. (2006) 'Conceptual framework and models for identifying and organizing usability impact factors of mobile phones,' *Proceedings of the 18th Australia Conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments*, Sydney, pp. 261–268.
- Drummond, J.S. and Themessl-Huber, M. (2007) 'The cyclical process of action research: the contribution of Gilles Deleuze,' *Action Research* 5 (4), pp. 430–448.
- Hornbæk, K. and Law, E.L.-C. (2007) 'Meta-analysis of correlations among usability measures,' *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, San Jose, CA, pp. 617–626.
- International Organization for Standardization (ISO) (1998) 'ISO 9241-11' (Guidance on usability); available at: [www.iso.org/iso/home.htm](http://www.iso.org/iso/home.htm).
- Isomursu, M., Tähti, M., Väinämö, S., and Kuutti, K. (2007) 'Experimental evaluation of five methods for collecting emotions in field settings with mobile applications,' *International Journal of Human-Computer Studies* 65 (4), pp. 404–418.
- Kirakowski, J. and Corbett, M. (1993). 'SUMI: the Software usability measurement inventory,' *British Journal of Educational Technology* 24 (3), pp. 210–212.
- Kjeldskov, J. and Stage, J. (2004) 'New techniques for usability evaluation of mobile systems,' *International Journal of Human-Computer Studies* 60 (5/6), pp. 599–620.
- MacKenzie, I.S. and Soukoreff, R.W. (2003) 'Phrase sets for evaluating text entry techniques,' *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, Ft. Lauderdale, FL, pp. 754–55.
- Macleod, M. and Rengger, R. (1993) 'The development of DRUM: a software tool for video-assisted usability evaluation,' *Proceedings of the HCI'93 Conference on People and Computers VIII*, Loughborough, UK, pp. 293–309; available at: <http://www.nigelbevan.com/papers/drum93.pdf>.
- Scapin, D. and Law, E.L.-C., eds. (2007) *Review, Report and Refine Usability Evaluation Methods (R<sup>3</sup>UEMs)*. Athens: COST294-MAUSE 3<sup>rd</sup> International Workshop.
- Stafford, T.F. and Gillenson, M.L. (2003) 'Mobile commerce: what it is and what it could be,' *Communications of the ACM* 46 (12), pp. 33–34.
- Steinar, K. and Ljungberg, F. (1999) 'Mobile informatics: innovation of IT use in mobile settings: IRIS'21 workshop report,' *ACM SIGCHI Bulletin* 31 (1), pp. 29–34.
- Zhang, D. and Adipa, B. (2005) 'Challenges, methodologies, and issues in the usability testing of mobile applications,' *International Journal of Human-Computer Studies* 18 (3), pp. 293–308.

# Mobility, Space, and Deserts: Urban Areas in Need of Precipitation

**Bob Jeffery**, *School of English, Sociology, Politics and Contemporary History, University of Salford*

## Abstract

This article reviews contemporary debates around the supposed existence of urban areas where the local population is unable to access a balanced variety of foodstuffs (so called "food deserts"), and the relationship of these phenomena to concerns regarding social and spatial exclusion. While the concept is disputed, evidence suggests that urban food (or more broadly "service") deserts are the result of a number of contingent and variable social processes that serve to create problems of access for specific sub-populations. These include those already recognised as suffering social exclusion, those in relative poverty, the elderly, and certain ethnic minority populations. A number of theories for understanding the dynamics of socio-spatial exclusion are considered, and it is argued (following Shaw, 2006) that the most pertinent are those that describe the interplay of three elements (abilities, assets, and attitudes). The second half of the essay is concerned with relating the theory of "food deserts" to a research project that is currently being conducted in a case study site in the Greater Manchester area, which is aimed at exploring links between mobility, social inequality (or exclusion) and identity.

**Keywords:** Mobility, transport, space, urban form, 'food desert', social exclusion, access to services.

## Deserts as spatial exclusion

*There are deserts in every life, and the desert must be depicted if we are to give a fair and complete idea of the country*

—Andre Maurois (French biographer, novelist, and essayist, 1885–1967)

An important problem in understanding the relationship of mobility to socioeconomic inequity in cities relates to urban form itself. New Labour's Social Exclusion Unit (SEU) has recognised that exclusion is partly constituted through the accessibility (or lack thereof) of key services. The Unit suggests that

Poor transport contributes to social exclusion in two ways. First, it can stop people from participating in work, learning, health care, food shopping and other activities such as volunteering and community participation. Second, people in deprived communities suffer the worst effects of road traffic through pollution and pedestrian accidents (SEU, 2002: 3).

Focusing on the first of these issues, the fact that transport is required at all to deliver access to retail, employment and public services suggests that these services cannot be found in one's immediate locality. Thus we are led to the conclusion that an understanding of urban form (and its historical development) is necessary for the elaboration of a theory of mobility related socio economic inequality. In building its case, the SEU goes on to cite the National Travel Survey (1998/2000), which reported that 16% of those without cars found it difficult to access food shops and supermarkets (as compared with only 6% of the wider population (SEU, 2002: 10). The report highlights similar disparities in terms of accessing health care, and particularly, hospitals (*ibid.*: 9). In *Making the Connections: Final Report on Transport and Social Exclusion*, "participation in social, cultural and leisure activities" is also recognised as "very important to people's quality of life" (SEU, 2003: 16). In addition, the problem of

exclusion from these activities is constructed in terms of access, transport, and a presumed dearth of such activities and facilities in particular localities.

One term in particular, "food deserts," has embedded itself in urban planning and public policy to describe the localities that fail to sustain their own decentralised food retail outlets. As food deserts play a prominent role in discourses surrounding urban form, mobility and access, this paper will focus on the development and the empirical base of this particular discourse. However, towards the end of this piece, we will suggest that the concept can be broadened to that of a "service desert," encompassing food, health, banking, post offices, and more.

The term "food deserts" originated from a working group of the Low Income Project Team of the Nutrition Task Force (Beaumont et al., 1995; Cummins and Macintyre, 1999). They defined food deserts as "areas of relative exclusion where people experience physical and economic barriers to accessing healthy food" (Reisig and Hobbiss, 2000, in Shaw, 2006). While the terminology has worked its way down to local government level (see Manchester City Council, 2008), the actual existence of food deserts is an issue of some dispute (see Cummins and Macintyre, 2002), and indeed it appears that the evidence that has been collected to date offers a particularly mixed picture. Nevertheless, by close analysis of the existing data it may be possible to isolate particular social groups at risk of inhabiting a notional food desert, and of the kinds of processes relating to urban form, historical and economic development, and political governance that have forged these spatialities where access to healthy food is at issue. Finally, an attempt will be made to relate what is now known about food deserts to the case study site of *Oldsville*—the pseudonym of a neighbourhood in the Greater Manchester conurbation—where we are investigating broader issues regarding the relationships between mobility, urban form, and socioeconomic inequalities.

### **The origins of deserts**

In their study of the putative existence of food deserts in Glasgow, Cummins and Macintyre discussed the phenomena in the context of restructuring of the retail sector. They argue that

[d]uring the 1980s there was a radical re-organisation of retail space as food retailing firms became larger and their total numbers diminished. These firms began to compete by concentrating on capital investment in large superstores often situated on edge or out-of-town sites (Cummins and Macintyre, 1999: 546).

An increase in the size of supermarket stores was driven by the desire for efficiency savings. Towards the end of the 1980s, the number of city centre sites suitable for redevelopment had diminished and land values increased, further prompting retailers to move towards more peripheral locations or into purpose-built retail parks (*ibid.*). These sites provided ample land for the vast car parks that would be required to service superstore-size food retail outlets, as well as being "usually more accessible to highways" (Larsen and Gilliland, 2008: 2).

A longitudinal study by Guy (1996, in Cummins and Macintyre, 1999) on the impact of these changes in Cardiff concluded that an initial burst of superstore development in the mid-1980s was followed by a wave of closures in the small grocery store sector, suggesting that centralisation in the food retailing industry came at the cost of both the number and diversity of outlets. Guy's study is borne out by the work of the SEU, whose data from the Department of Health suggests the number of small shops fell by 40% nationally between 1986–1997 (SEU, 2002: 10). This in turn has had an impact on the spatial dispersion of food retail outlets, and thus we may argue, on the ability of differently placed populations to access these outlets.

In the analysis of their city of Glasgow case study site, Cummins and Macintyre (1999: 549) conclude that areas without shops are primarily found in rural or semi-rural locations. Also, as



we might expect, they note that the highest density of stores (including multiples, as superstores are referred to in planning jargon) are typically found in the most densely populated urban areas. (This suggests that Jacobs's (1994[1961]: 213–17) theses regarding population density and commercial vitality still hold true. Cummins and Macintyre also noted a concentration of "discount stores" (such as Aldi, Lidl, and Netto) in the poorer areas. This is significant, as the Social Exclusion Unit, citing the Consumers Association, argues that the cost of basic foodstuffs is greater in local shops than in supermarkets (SEU, 2003: 15–16). This would suggest that those unable to leave their neighbourhoods to access larger stores would have to pay a premium for their food. However, if "discounters" are clustered in deprived areas, this may counterbalance such an effect. The authors conclude that at the time of their study, deprived areas were more likely to host food outlets than comparatively affluent ones, which suggests an absence of food deserts (Cummins and Macintyre, 1999: 552). At the same time, they acknowledge that it is still to be proven whether such findings would be replicated at a lower level of spatial analysis (the postcode sector level) rather than that used in the study (postcode district level, *ibid.*). It is plausible that geographically specific and fragmented food deserts are submerged beneath the seeming ease of access at the broader level.

Larsen and Gilliland (2008) have carried out longitudinal Geographic Information Systems-based research on the suburbanisation of food retailing in a small Canadian city. In addition to the issues of the premium price paid by the clientele of small food retailers identified by the SEU, the authors also note that "[w]hile supermarkets also carry unhealthy foods (e.g. chips, soft drinks, and processed foods), these items are more readily available at neighbourhood convenience stores, which are less likely to offer items supportive of a healthy diet" (*ibid.*: 1–2). This somewhat broadens the notion of access-based-exclusion from food retailers per se to food retailers that stock a balance of affordable and healthy foodstuffs.

Their work also explicitly focuses on the impact of automobility (the automobile system; see Urry, 2007) on the spatial form of urban environments (Larsen and Gilliland, 2008: 2). A society of mass automobility stretches the "activity space"—the space in which daily activities are contained—of its inhabitants (Axhausen, 2007; see also Jacobs, 1994[1961] and Hanson, 1995). It is productive of forced mobility (see Sager, 2006: 472–473), even for those with limited transport options. As Larsen and Gilliland argue,

Investment in highway infrastructure, particularly in the period immediately following WWII, cemented the dominance of the automobile as the preferred mode of transit. This form of privatized mobility, however, was not democratically distributed; only those with sufficient wealth could afford the luxury of auto-mobility (Larsen and Gilliland, 2008: 2).

It is necessary to note here, that while there are important differences between the phenomenon of suburbanisation or urban sprawl—as evidenced in North America, the United Kingdom, and mainland Europe—we might suggest that the UK is rather closer to the USA than its continental neighbours. Both of these countries ardently subscribed to the reactionary anti-urbanism of Howard's Garden Cities, leading "eventually to the 'New Towns' programme of planned suburbanisation" (Pichler-Milanovic, 2007: 104). On the continent, meanwhile, the ideals of Le Corbusier took hold, with his stress upon remodelling the inner city and using modern technology to increase urban densities (*ibid.*: 104–05).<sup>1</sup> This emphasises how examples of the same phenomena (food deserts and urban form) will vary across societies, and implies that we must exercise caution in not overstretching our comparative material.

---

<sup>1</sup> As Hanson (1995: 28) asserts, "The European metropolis, though now experiencing decentralisation, retains a more tightly agglomerated spatial structure, and the historical central city continues to dominate its immediate urban region." (See also Giuliano, in Hoyle and Knowles, 1998: 127–30.)

Nevertheless, although suburbanisation to some extent predates the rise of the automobile<sup>2</sup>, it is clear that the shift to mass automobility and the “predict and provide” planning models that established the rationale for the huge investment in automobile infrastructures (such as motorways; see Goodwin, 1999) have had a massive impact upon the spatial forms extant in the UK, which though not of the same degree, certainly parallel those found in North America.

As Larsen and Gilliland detail, at first retailers, offices and other commercial functions remained in the city while residents suburbanised, but this came to an end in the 1970s, when these businesses began to move closer to their suburban base. This, the authors contend—although driven by the factors considered above—was facilitated through “sophisticated location analysis” in order to provide “the most rational spatial arrangement of stores for maximal profits” (Larsen and Gilliland, 2008: 2).

Following this elaboration of the historical context, the authors present a review of the existing literature on food deserts and conclude that the picture presented by empirical data is mixed for the United States, Canada, and the UK. In the US there seems to be a relationship between food deserts and ethnic minority populations (a theme returned to below), while in Canada, studies have shown poor inner-city neighbourhoods (such as our UK case study site, *Oldsville*) actually poses the best supermarket access (*ibid.*: 3).

What is of particular interest with regards to the study of London, Ontario, is the method of geographical analysis used—namely one based on “a network . . . approach using the city’s circulation system (e.g., streets, footpaths) and public transit routes (bus lines) in calculating the distance one can travel,” as opposed to using a “circular “buffer” to identify all sites located within a pre-determined, straight-line distance (or ring) from origin” (*ibid.*: 4). In short, the methods of analysis are much more sensitive to the particularities of urban form, and to the two modes of transport (walking and buses) most accessible to socially excluded populations.

The results of this study were that supermarket access has diminished over time and that “the average proportion of census tract population with easy supermarket access in 1961 (45.2 per cent) was more than twice the level for 2005 (18.3 per cent)” (*ibid.*: 8). They further demonstrate that the most socioeconomically deprived neighbourhoods have the poorest access to supermarkets by foot (measured as a 1-kilometre walk)—23.8%, as compared to the least deprived neighbourhoods, where 30.3% have such access. Conversely, the least deprived neighbourhoods have 81.6% access by public transport (bus), and this was lower than the most deprived areas at 86.4% (*ibid.*: 8, 11).

So on the one hand, deprived neighbourhoods (primarily clustered around the inner city) are further from supermarkets on foot. But on the other, by virtue of being located in the inner city, they are in the best possible position to utilise bus routes that converge on the city centre in a radial fashion. This has two implications for our research focus on social exclusion. Firstly, while inhabitants of more affluent areas are more likely to have access to a car (SEU, 2003: 29) and be located closer to a supermarket, those few that do not may be considered to be living in a “food desert,” and one exacerbated by fewer public transport routes flowing through it. On the other hand, those in inner-city deprived neighbourhoods (only around 50% of those in the lowest income decile have access to a car in England; *ibid.*) may have access to supermarkets by public transport, but this indeed may be their only option. Thus the affordability of fares to those who have the least disposable income must be considered of vital importance.

In summing up, the authors compare their work with an analysis of the much larger Canadian city of Montreal, which found that most areas had good supermarket access and concluded

---

<sup>2</sup> Pooley and Turnbull (1998: 112) contend that counter-urbanisation occurred from the nineteenth century onwards, whereas it is usually seen as a twentieth-century phenomena (see also Shapeley, 2007).

that food deserts did not in fact exist. This leads Larsen and Gilliland to suggest that the presence of food deserts in London, Ontario, may be a function of differing populations and urban forms: "Montreal is much larger than London in terms of total population, at 1.8 million versus 350,000 inhabitants, and it also has a much higher population density: 847 versus 170 per square kilometre" (Larsen and Gilliland, 2008: 13)). We can infer from this that lack of retail provision is intimately connected to lower spatial densities, suburbanisation, and in Jacobs's (1994[1961]: 391) terms, a lack of city vitality. This assists us in interrogating the urban form of our UK case study site when attempting to understand spatial exclusion.

### **Deserts of the real (or real deserts?)**

According to Cummins and Macintyre, the existence of food deserts is open to dispute. They suggest that uncritical citations of the scant evidence that exists have been widespread: "The Social Exclusion Unit's report gave no supporting evidence for the assertion that some urban areas of the United Kingdom had become food deserts" (Cummins and Macintyre, 2002: 437). They cite both their own research (the Glasgow study) and that of Mooney (1990), which suggest that large multiple stores were more likely to be found in deprived areas, and that both "healthy" and "unhealthy" baskets of shopping were cheaper in deprived areas than more affluent ones. They conclude by asserting that a "food desert" is a "factoid": "assumptions or speculations reported and repeated so often that they are popularly considered true: they are simulated or imagined facts" (2002: 436; see also Wrigley et al., 2004). How then are we to reconcile the mixed evidence we have reviewed so far?

Shaw (2006) presents us with a way forward in his attempt to develop a classificatory schema of differing food desert types, and it is to this work that we now turn. Firstly, Shaw begins by asking us to think about the kinds of barriers that may prevent various social groups from accessing food retail outlets and healthy dietary options. He begins by arguing that we need to consider issues as diverse as "weight carrying capacity,"<sup>3</sup> the requirement of adequate refrigerated space in which to store perishable foodstuffs, and the anticipation of lengthy queues (*ibid.*: 232) in order to understand perceptions of distance and convenience. He concludes this review by arguing that "the factors debarring access to a healthy diet are multiple and extend well beyond strictly spatial or geographical aspects" (*ibid.*: 233). This is a finding of critical importance, as it takes us beyond the physical structures of cities to the subjective social constructions of groups and individuals.

Shaw proceeds to describe how such geographical aspects intertwine with socioeconomic status, the distribution of foodstuffs, and individual perceptions by way of an enumeration of different categories of "food deserts." In the first of these, *food access problems in more affluent suburbs*, he discusses how not all residents may have access to a car and even households that do have the potential to lose such access through "accident, theft, disqualification, or failing health," and that access may be particularly problematic for pensioners (*ibid.*: 238). In the second, *attitudinally and demographically based grocery access problems*, we find a different set of processes at work. In his Leeds case study site, Shaw describes the problem of student-dominated areas, where tastes may dictate a shift to convenience foods to the detriment of the choices available to other social groups. In addition to this, grocery provision may become seasonal, in line with the university calendar, thus creating temporal food deserts (*ibid.*; see also Hubbard, 2008 on the wider effects of "studentification"). In the second attitudinal example, Shaw points to the South Asian-dominated areas of Southwest Headingley and Beeston, where "[b]utchers serving the white pensioner population are closing contemporaneously with new Halal butchers opening" (*ibid.*). This implies that while on the surface these areas may have many food retail outlets, the perceived "foodscape" may vary enormously on the basis of ethno-cultural dispositions. ("Many white working-class pensioners want potatoes, they don't like seeing 'exotic' fruit and vegetables"—interview respondent, in *ibid.*).

---

<sup>3</sup> "Food is relatively heavy; consumers can only carry so much, and some of their weight carrying capacity is already spoken for by essential non-food items (e.g. household cleaning goods or pet food)" (Shaw, 2006: 232).

Staying with Leeds' South Asian population, a Bangladeshi community group interviewed for the purposes of the research commented: "Some Bangladeshi women lack confidence, and there is a problem with muggings by youths; few have cars as earnings are low. The most the women will walk to the shops is to the corner shop, or maybe five or ten minutes' walk" (*ibid.*: 238–39). The inference is that fear of crime combined with the cultural traditions of the community serve to isolate some individuals to a greater extent than would be suggested by geography alone.

With *time-based constraints to healthy food*, Shaw discusses the cash-rich but time-poor, who, despite their relative affluence, consume large quantities of pre-packaged and processed foods. The fourth category of food desert is defined by a *lack of cooking skills*, which may be linked to low socioeconomic status and wider changes in the food retail market:

Once households are deskilled in cooking expertise, their food purchasing may shift from fresh fruit, meat, and vegetables towards pre-packed ready meals, which will tend to be offered within a greater range and at a lower cost in a large supermarket than in a local grocery store. A vicious circle may ensue with the closure of local retailers precipitating further cookery deskilling and continued shift of trade to distant supermarkets, leaving behind altogether those of limited mobility and means (Shaw, 2006: 240).

Finally, Shaw elaborates on *wealth-based constraints to healthy eating and other constraints on a healthy diet*. In the first of these, ethno-cultural tastes are again at issue, as research by the author in Scunthorpe has identified that foods preferred by immigrant communities, such as South and East Asians, may cost more "due to limited economies of scale and the import route of these foods through agencies located in cities such as Manchester, for historical reasons" (*ibid.*). In addition to this, we have the various forms of cost-based exclusion relating to transport. In the final category, the author considers various miscellaneous issues, such as inability to prepare certain foods due to arthritis in aged populations and the lack of cooking facilities available to asylum seekers (*ibid.*: 240–41).

In attempting to conceptually reduce the diversity of food deserts that may be found in urban and rural environments, Shaw posits a threefold classification of: "ability," "asset," and "attitude," which in combination account for the variance between different types of food deserts as they are experienced by different social groups. Here we find a point of confluence with other theories of mobility-related exclusion, in particular with those of Church et al. (2000), Kaufmann et al. (2004), Kaufmann and Flamm (2006), and Urry (2007).

Shaw (2006: 241–42) defines "ability" as problems of "physical access," "asset" as financial resources, and "attitude" as the more intangible layers of knowledge, familiarity, and sociocultural dispositions. We can compare this to Church et al.'s notion of the constituting elements of transport related exclusion, while recognising how these elements are reduced in Shaw's understanding: economic- and time-based exclusion may be subsumed under "asset"; geographical-, physical- and facility-based exclusion under "ability"; and fear-based exclusion under "attitude" (see Church et al., 2000: 198–200). Similarly, Shaw's schema can be interpreted as capturing the elements described in Urry's (2007: 194) notion of "network capital."

Network capital as a conceptual device aims to encapsulate what Urry considers to be the four elements constitutive of the notion of "access": economic, physical, organisational, and temporal. Network capital in turn is seen as having eight elements which "*in their combination* produce a distinct stratification order" (*ibid.*: 197). The elements incorporate issues as diverse as the ability to read maps and timetables to being able to access safe and convenient meeting places (*ibid.*: 198). Finally, a third conceptual schema is advanced by Kauffman et al. (2004) and Kaufmann and Flamm (2006) and encapsulated within the idea of "motility" (or potential

for movement). Kaufmann et al. (2004: 750) identify three constituent features: *access* (place, time and other contextual constraints, networks, and dynamics within territories); *competence* (physical ability, acquired skills, and organisational skills), and *appropriation* (strategies, motives, values, and habits). What is important about Kaufmann et al.'s (2004) analytical schema is not so much access (found in Church et al., 2000) or competences (a key focus of Urry's [2007] work), but with appropriation, which speaks of subjective rational and non-rational motivations in mobility-related decision making. In this sense, the concept presented by Kaufmann et al. (2004) shares a similarity with Shaw's notion that "attitudes" exert a practical influence upon the mobility patterns of individuals, and their ability to access particular sites.

What is of vital importance to note is that whatever classificatory schema we adopt, the various options presented are to some degree interchangeable. We must therefore *elaborate in full* what these various generalisations may mean in light of *individual narratives* and how they are played out in *particular spatial contexts*. It is clear from the evidence we have reviewed thus far that food deserts are complex entities, consisting of urban form, transport networks, and actors with varying socioeconomic and sociocultural resources. It seems plausible that neighbourhoods with very similar geographies or very similar social groupings may have widely diverging experiences of access-related exclusion.

### **In search of deserts and oases (method)**

To evaluate the efficacy of the "food desert" construct, as well as the various theories of mobility-related exclusion, we will now attempt to specify the ways in which they can be applied to our case study site of *Oldsville*, located some 2 kilometres from the Greater Manchester regional centre. According to the New Labour government's *English Indices of Deprivation 2007*, the area of *Oldsville* is one the 100 most deprived Super Output Area's in the country (broadly corresponding to ward-level, out of a total of 7,193 middle-layer Super Output Areas in England; see Department for Communities and Local Government, 2008: 113).<sup>4</sup> Significantly, the ward was not placed in the worst 100 at the time of the prior *Indices* study by the Office of the Deputy Prime Minister (2004), although this may be the result of changes in reporting standards. The indices are a compound score that is a function of seven principle components: income deprivation; employment deprivation; health deprivation and disability; education, skills, and training deprivation; barriers to housing and services; crime; and living environment deprivation (Department for Communities and Local Government, 2008: 33).

*Oldsville* has a long history of industrial employment, as well providing accommodation for nearby dockyards and warehouses. From the mid to late-twentieth century *Oldsville* was also a site of mass slum clearances (Shapely, 2007: 98). In addition to the untold disruption of existing social networks in the inner cities, the former slum dwellers who moved onto overspill estates found themselves with longer, and crucially, more expensive commutes to places of employment, along with a lack of community institutions (pubs, small businesses) on the new estates (*ibid.*: 69). In the 1970s, both the Docks and the wider area went into terminal decline (Struthers, 2003: 11). This was linked to wider industrial restructuring within Greater Manchester. In between 1959 and 2000, the conurbation's manufacturing workforce declined from well over half to less than one-fifth (Peck and Ward, 2002: 1).

While Greater Manchester has made something of a recovery since the 1990s, this has been mainly service-led and has caused a wholesale restructuring of local labour markets (see Giordano and Twomey, in Peck and Ward, 2002: 50–75). This is of particular relevance to *Oldsville*, as this shift has been marked by changes in the spatial distribution of labour within the conurbation, with sites on the periphery of the conurbation holding greater attraction for employers: "By the mid-1970s the focus of employment growth in the North West region had

---

<sup>4</sup> Middle-layer Super Output Areas discussed in the report were matched with local areas utilising the "Neighbourhood Statistics" webpage of the Office for National Statistics website. Available at: <http://www.neighbourhood.statistics.gov.uk/dissemination>.

shifted to outlying towns like Warrington and Wigan, both of which were well served by the regional motorway network" (*ibid.*: 51).

By the 1980s *Oldsville* had been utterly transformed. Its traditional terraces had largely been demolished in the 1960s and 1970s and replaced by "system built high and low rise housing stock" (Walklate, 1998: 553). Of particular interest to our foci is Walklate's contention that as unemployment worsened throughout the 1980s: "[L]ocal road-building policies literally cut the area off from its surroundings on all sides with many of its local amenities being demolished in the process (*ibid.*: 554).

This last point is particularly pertinent for our purposes, as it speaks to the ways in which the dynamism and potential cross-use of the *Oldsville* neighbourhood have been curtailed by ill-considered planning policies. As Jacobs (1994[1961]: 275) has suggested in her seminal *The Death and Life of Great American Cities*, "[s]ome borders halt cross-use from both sides. Open railroad tracks or expressways or water barriers are common examples."<sup>5</sup> Such examples of semi-permeable barriers surround *Oldsville* on all sides, from the four-lane arterial route that marks the northern frontier of the neighbourhood and links to the city centre bypass, to the river, canals, and other major trunk roads that surround the neighbourhood on all sides.

In addition to the physical impact of carriageways, *Oldsville* is marked by both an unusually low population density (for an inner-city neighbourhood) and by numerous "brownfield" sites awaiting redevelopment. In the first instance, the area has a population density of 16 people per hectare (Salford City Council, 2009). In Jacobs's (1994[1961]: 46) terms, this is far too low to create city vitality or to put enough "eyes on the street" to create the kind of "natural surveillance" that is so vital giving people the confidence to walk the streets feeling safe from harassment. The first issue of residential densities brings us back to Larsen and Gilliland's contention that supermarket access may be related to urban density and form. While *Oldsville*'s number of residents per square kilometre (2,000) is greatly in excess of both Montreal (847) and London, Ontario (170), this may primarily be due to the fact that we are comparing a single city district with larger conurbations, whose overall densities will be dragged down by swathes of suburban residential developments. Nevertheless, as of the 2001 Census, *Oldsville* was the least densely populated electoral ward in the city, and well below the sufficient levels suggested by Jacobs (*ibid.*: 217).

The second issue, the lack of "natural surveillance," is apt to make people feel insecure as "dead spaces" (*ibid.*: 277) become magnets for criminality and antisocial behaviour. While *Oldsville* is no longer considered as dangerous an area to inhabit as it was at the time of Walklate's (1998) survey, clearly we must recognise how *fear-based* exclusion (in Church et al.'s [2000] terms) may be a factor in determining the accessibility of various services (including access to food). This will of course vary across groups and individuals. Preliminary analysis of the interview data in *Oldsville* highlights the case of an elderly woman who will not "leave the house on foot at night." However, for younger residents this does not seem to be an issue.

To conclude this analysis of our case study site, we will present the findings of two small-scale studies (n = 67 and n = 156) on access to amenities, conducted by a local commercial developer (who will remain anonymous for the purposes of this study).<sup>6</sup> The survey sampled the full range of geographical locations across the *Oldsville* estate. The first survey revealed that 68% of respondents were very dissatisfied, and another 25% dissatisfied with the current range of shops in the neighbourhood. Only 3% were satisfied. It also revealed that 58% used

---

<sup>5</sup> See also Graham and Marvin (2001: 120–21) on "residualising the street" and SEU (2003: 17) for an acknowledgment of this issue by the government's Social Exclusion Unit

<sup>6</sup> The survey was conducted by a team commissioned by a local commercial developer, a representative of which provided the data to the author. The developer remains anonymous both to protect the anonymity of the neighbourhood in question, as well as to protect the commercial interests of said developer.

a local supermarket that is commonly recognised as being amongst the higher-end of retailers in terms of expense. From extensive participant observation, it would appear that the higher social class of many of the patrons of this retailer, as compared to the community in which it is situated, is related to the car access afforded by its location (adjacent to the main trunk route leading to the regional centre, and a short distance from a major urban regeneration development). The argument here would be that the supermarket is so situated because of the ease of access it affords to highly mobile individuals, whereas those of low mobility in the socially deprived *Oldsville* neighbourhood have no access to lower cost alternatives.<sup>7</sup>

The survey also found that 52% of those interviewed use a car to complete their weekly shop (this broadly correlates with the percentage of car access amongst the poorest decile of the wider population of England; SEU, 2003: 29), while 22% walk. The remainder use taxis, bicycles, a wheelchair, or public transport. Clearly, we must be cautious about the representativeness of a survey conducted using a very small sample ( $n = 67$ ), and for commercial purposes. Nevertheless, the survey does offer an impression of the issues around spatial access that exist in *Oldsville*. The second survey ( $n = 156$ ) corroborates these impressions, recording that around 50% of individuals surveyed commonly shopped at the aforementioned supermarket, while the remainder accessed a narrower range of small retailers, which we would expect to be both more expensive and to offer a narrow range of healthy foods. When asked, "What shops in other areas do you use?," the survey recorded that around half accessed a range of shops at a local precinct some 2 kilometres from *Oldsville*. It was only the remainder that left the area to access superstores (much larger than the supermarket in *Oldsville*) with a greater variety of foodstuffs at lower costs. Finally, when asked what services they would like to see in their local area, 106 individuals called for a post office and 80 requested an a(nother) supermarket. This offers us a clear indication that some individuals in the *Oldsville* neighbourhood consider their area to be lacking in services, and indeed that a comparatively large percentage do not access the range of services found at a wider urban level. The fact that a range of services were felt to be lacking in the case study locality (including hairdressers) justifies expanding the concept of "food desert" to "service desert."<sup>8</sup>

## Emerging themes

In this penultimate section, I propose to briefly discuss particular themes that have emerged from participant observation and exploratory interviewing with residents of *Oldsville*, before suggesting directions for future research and sketching out how the theme of service deserts relates to the wider problems of space, mobility, and socioeconomic inequality.

The first issue facing residents relates to the (lack of) public transport provision in the local area. While the neighbourhood is located close the regional centre of the conurbation, bus services are infrequent and highly unreliable, and several interviewees have commented on this (one even keeps a file on the failings of a particular bus route).<sup>9</sup> Related research is aimed at unravelling the local transport policies that structure the current configuration of provision

---

<sup>7</sup> There is an interesting parallel between the *Oldsville* supermarket, which appears to be targeted at commuting non-residents, and the views of local residents concerning the impact of a new store (an intervention study) in the East Leeds food desert (Wrigley et al., 2004). While the new supermarket did improve access to a range of balanced and affordable foods, the impact felt was modest (*ibid.*: 134). Further, its presence was even resented by some local residents, who perceived the store as targeted towards more affluent "outsiders." This was expressed by one focus group participant in the following way: "You see people with a full basket of shopping and you think that must be a hundred quid. I could never spend that much" (*ibid.*: 130). The authors concluded that the new store, positioned as a mainstream and non-discount outlet was merely another factor to be coped with in the struggle to eat an adequate diet on a low income (*ibid.*: 134).

<sup>8</sup> One constituent of "service deserts," lack of banking access, receives treatment by the New Economics Foundation in their report, *Ghost Town Britain II: Death on the High Street*. Here they report that "[m]ore than 800 communities across Britain have already lost all local banking presence" (2003: 15).

<sup>9</sup> The respondents interviewed by Pooley et al. for their survey of mobility in Salford also reported problems concerning bus services, with poor reliability and the necessity of numerous connections when making cross-conurbation journeys (2005: 35)

within Greater Manchester.<sup>10</sup> At the same time, other respondents have commented on how bus services had been withdrawn from the local area in the past due to crime and “antisocial behaviour” (see Granville and Campbell-Jack, 2005 on antisocial behaviour on buses).

Secondly, a high degree of reliance upon taxis has been observed in this area, and this seems to be the preferred mode of transport from some individuals, particularly those with caring responsibilities. This is consistent with observations from Shaw’s Leeds study (2006: 236). Larsen and Gilliland accurately diagnose the exclusionary mechanisms implied in such reliance:

When walking is not feasible and transit access is not available, residents may use a taxi for transportation, but taxi fare adds a significant cost to the household budget which can be troublesome for persons of low income or fixed incomes. Furthermore, taxi dependency would also necessitate making fewer trips to the supermarket to avoid multiple fares; and, in turn, less frequent trips would be detrimental to the household’s ability to take full advantage of time-sensitive sales, to keep fresh produce in the home, and to minimise the total financial outlay for anyone shopping trip (Larsen and Gilliland, 2008: 12).

Other issues relate to road safety, problems of accessing childcare facilities, open spaces, perceptions of crime and safety, and access to employment. As the process of conducting qualitative research interviews continues, we will be in a better position to try to understand service deserts as one aspect of mobility-related exclusion, as well as the various contingencies that lead to differing outcomes across groups and individuals. At this early stage, it is nevertheless possible to imagine how the classificatory schemas of Shaw and Kaufmann in particular, may be useful in accounting for the various complexities of spatial exclusion. These include having a driving licence or the money for a fare (*asset, competence*); being on a bus route or living near a post office (*ability, access*); the desire for certain foodstuffs; and the wish to avoid certain areas at night (*attitude, appropriation*) (Shaw, 2006; Kaufmann et al., 2004).

## Conclusions

In conclusion, there is convincing evidence that, contrary to the assertions of Cummins and Macintyre, food (or service) deserts do exist, at least for some people, some of the time. This undoubtedly imposes extra layers of obstacles, inconveniences, and daily difficulties upon various social groups, but most especially those populations already recognised as being socially excluded by dint of their status (such as single mothers; see Chlond and Ottmann, 2007) or due to their locality. (Clearly, the two are not unrelated; see Byrne, 1999: 87–90 on the spatial clustering of single parents.) These issues are intertwined with wider processes of social exclusion related to a lack of economic resources (and often temporal resources too), and to the historic neglect of particular localities—all compounded by the incredible influence of the car over urban form as *the* mode of mass privatised transport.

This study of the problem of food deserts is only one part of a wider study of mobility, inequality, and identities. In much of the literature, “food deserts” have been falsely assumed to be a purely spatial, that is a geographical, problem (see Cummins and Macintyre, 1999). It is clear that there are also important socio-demographic factors at play, as well as the intersection of uniquely placed social actors with their own particular narratives. At the macro level, various social theorists have expressed concern about the power relations of mobility in

---

<sup>10</sup> A preliminary review of the literature regarding transport policy in Greater Manchester has revealed a huge decline in the number of passengers carried by public transport over the twentieth century (Pooley et al., 2005); that the average bus journey has increased in length (Knowles, R. et al., 1991: 18); that the average journey to work has more than doubled since the 1960s (Pooley et al., 2005: 20); and that 33% of households across the conurbation do not have access to a car (Greater Manchester Passenger Transport Executive, 2006). The reduction in public transport services is also heavily connected to privatisation and deregulation (van Goeverden et al., 2006: 23).



contemporary capitalist societies (see Bauman, 2000, Castells, 2000, Urry, 2007, and even the discussion of a “kinetic elite” by Koolhaas, 1996). The formulation is expressed by Morley in the following way: “Mobility is increasingly seen as a social good and immobility increasingly acquires, by contrast, the connotation of defeat, of failure, and of being left behind” (2000: 202). The aim of the broader research project is to tease out these connections and to test the hypothesis that localities that are objectively isolated (at least in relative terms) produce identities that are *bounded to place* (see Kintrea et al., 2008 on “territoriality as a cultural expectation”). In doing so, we must recognise that space does not explain everything, and must also look to social networks to understand dispositions towards wider forms of mobility than have not been considered here (residential mobility as opposed to daily mobility/commuting). Such work would consider how social networks (MacDonald et al., 2005), class-based identifications (Southerton, 2002; Tyler, 2008), and the twin tropes of cosmopolitanism and localism (Thompson and Taylor, 2005) militate for or against the spatial and socioeconomic structures that influence geographical mobility. In so doing, it will contribute to the growing body of work across a range of disciplines (sociology, social policy, human geography, and cultural studies) that aims to specify the relationship of transport and mobility to wider processes of social exclusion and to campaign for their mitigation.

## References

- Bauman, Z. (2000) *Liquid Modernity* Cambridge: Polity Press.
- Beaumont, J. et al. (1995) *Report from the Policy Sub-group to the Nutrition Task Force Low Income Project Team of the Department of Health*. Radlett: Institute of Grocery Distribution.
- Byrne, D. (1999) *Social Exclusion (Issues in Society)*. Buckingham: Open University Press.
- Castells, M. (2000) *The Rise of the Network Society*. Oxford: Blackwell Publishers.
- Church, A. et al. (2000) ‘Transport and social exclusion in London,’ *Transport Policy* 7, pp. 195–205.
- Chlond, B. and Ottmann, P. (2007) ‘The mobility behaviour of single parents and their activities outside the home,’ *The German Journal of Urban Studies* 46 (2). Available at: <http://www.difu.de/index.shtml?/publikationen/dfk/en/>.
- Cummins, S. and Macintyre, S. (1999) ‘The location of food stores in urban areas: a case study in Glasgow,’ *British Food Journal* 101 (7), pp. 545–53.
- Cummins, S. and Macintyre, S. (2002) “‘Food deserts’: evidence and assumption in health policy debate,’ *British Medical Journal* 325, pp. 436–38.
- Department for Communities and Local Government (2008) *The English Indices of Deprivation 2007*. Available at: <http://www.communities.gov.uk/documents/communities/pdf/733520.pdf>.
- Giordano, B. and Twomey, L. (2002) ‘Economic transitions: restructuring local labour markets,’ in Peck, J. and Ward, K., *City of Revolution: Restructuring Manchester*. Manchester: Manchester University Press.
- Giuliano, G. (1998) ‘Urban travel patterns,’ in Hoyles, B., and Knowles, R., *Modern Transport Geography, 2<sup>nd</sup> Edition*. Chichester: John Wiley & Sons Ltd.
- Goodwin, P. (1999) ‘Transformation of transport policy in Great Britain,’ *Transportation Research, Part A* 33, pp. 655–69.
- Graham, S. and Marvin, S. (2001) *Splintering Urbanism: Networked Infrastructures, Technological Mobilities, and the Urban Condition*. London and New York: Routledge.
- Granville, S. and Campbell-Jack, D. (2005) *Anti-social Behaviour on Buses*. The Scottish Government Development Department Research Programme, Research Findings No. 196. Available at: <http://www.scotland.gov.uk/Resource/Doc/37428/0029528.pdf> accessed 06/01/09)
- Greater Manchester Passenger Transport Executive (2006) *Transport Plan 2 (2006-2011)*. Available at: <http://www.gmltp.co.uk/pdfs/2006/GMLTP2.pdf>.

- Guy, C.M. (1996) 'Corporate strategies in food retailing and their local impacts: a case study of Cardiff,' *Environment and Planning A* 28, pp. 1575–1602.
- Hanson, S. (eds.) (1995) *The Geography of Urban Transportation*. New York and London: The Guildford Press.
- Hubbard, P. (2008) 'Regulating the social impacts of studentification: a Loughborough case study,' *Environment and Planning A* 40, pp. 323–41.
- Jacobs, J. (1994[1961]) *The Death and Life of Great American Cities*. Harmondsworth: Penguin Books.
- Kaufmann, V. et al. (2004) 'Motility: mobility as capital,' *International Journal of Urban and Regional Research* 28 (4), pp. 754–56.
- Kaufmann, V. and Flamm, M. (2006) 'Operationalising the concept of motility: qualitative study,' *Mobilities* 1 (2), pp. 167–89.
- Kintrea, K. et al. (2008) *Young People and Territoriality in British Cities*. Joseph Rowntree Foundation. Available at: <http://www.jrf.org.uk/bookshop/eBooks/2278-young-people-territoriality.pdf>.
- Knowles, R. et al. (1991) 'Recent transport developments in Greater Manchester,' *Manchester Geographer* 12, pp. 2–24.
- Koolhaas, R. (1996) *S, M, L, XL*. New York: Monacelli Press.
- Larsen, K. and Gilliland, J. (2008) 'Mapping the evolution of "food deserts" in a Canadian city: supermarket accessibility in London, Ontario, 1961–2005,' *International Journal of Health Geographies* 7, pp. 1–16. Available at: <http://www.ij-healthgeographics.com/content/pdf/1476-072X-7-16.pdf>.
- MacDonald, R. et al. (2005) 'Growing up in poor neighbourhoods: the significance of class and place in the extended transitions of "socially excluded" young adults,' *Sociology* 39 (5), pp. 873–91.
- Manchester City Council (2008) *Communities and Neighbourhoods Overview and Scrutiny: Access to Healthy and Sustainable Food*. Available at: [http://www.manchester.gov.uk/downloads/8c\\_Food\\_Deserts\\_1\\_.pdf](http://www.manchester.gov.uk/downloads/8c_Food_Deserts_1_.pdf).
- Mooney, C. (1990) 'Cost and availability of healthy food choices in a London health district,' *Journal of Human Nutrition and Dietetics* 3 (2), pp. 111–20.
- Morley, D. (2000) *Home Territories: Media, Mobility, Identity*. London: Routledge.
- New Economics Foundation (2003) *Ghost Town Britain II: Death on the High Street (How Britain's local economies are losing ground and fighting back)*. Available at: <http://www.neweconomics.org/gen/uploads/2hk0dtqzv0run55afsofu4515122003114309.pdf>.
- Office of the Deputy Prime Minister (2004) *The English Indices of Deprivation 2004 (Revised)*. Available at: <http://www.communities.gov.uk/documents/communities/pdf/131209.pdf>.
- Pichler-Milanovic, N. (2007) 'European urban sprawl: sustainability, cultures of (anti) urbanism and "hybrid cityscapes,' *Dela* 27. Available at: [http://www.ff.uni-lj.si/oddelki/geo/publikacije/dela/files/Dela\\_27/06\\_pichler.pdf](http://www.ff.uni-lj.si/oddelki/geo/publikacije/dela/files/Dela_27/06_pichler.pdf).
- Pooley, C. and Turnbull, J. (1998) *Migration and Mobility in Britain since the 18<sup>th</sup> Century*. London: UCL Press.
- Pooley, C. et al. (2005) *Travelling Around Town: Everyday Mobility in Manchester and Salford since the 1940s*. Cartographic Unit, Department of Geography, Lancaster University.
- Reisig, V. and Hobbiss, A. (2000) 'Food deserts and how to tackle them: a study of one city's approach,' *Health Education Journal* 59, pp. 137–49.
- Sager, T. (2006) 'Freedom as mobility: implications of the distinction between actual and potential travelling,' *Mobilities* 1 (3), pp. 465–88.
- Salford City Council, (2009) 'Salford ward profiles.' Available at: <http://www.salford.gov.uk/living/planning/planning-research/salfordwardprofiles.htm>.

- Shapely, P. (2007) *The Politics of Housing: Power, Consumers and Urban Culture* Manchester: Manchester University Press.
- Shaw, H. (2006) 'Food deserts: towards the development of a classification,' *Geografiska Annaler: Series B, Human Geography* 88 (2), pp. 231–47.
- Social Exclusion Unit (2002) *Making the Connections: Interim Findings from the Social Exclusion Unit*. Cabinet Office. Available at:  
[http://www.cabinetoffice.gov.uk/media/cabinetoffice/social\\_exclusion\\_task\\_force/assets/publications\\_1997\\_to\\_2006/making\\_transport\\_interim.pdf](http://www.cabinetoffice.gov.uk/media/cabinetoffice/social_exclusion_task_force/assets/publications_1997_to_2006/making_transport_interim.pdf).
- Social Exclusion Unit (2003) *Making the Connections: Final Report on Transport and Social Exclusion*. Cabinet Office. Available at:  
[http://www.cabinetoffice.gov.uk/media/cabinetoffice/social\\_exclusion\\_task\\_force/assets/publications\\_1997\\_to\\_2006/making\\_transport\\_summary\\_2003.pdf](http://www.cabinetoffice.gov.uk/media/cabinetoffice/social_exclusion_task_force/assets/publications_1997_to_2006/making_transport_summary_2003.pdf).
- Southerton, D. (2002) 'Boundaries of "us" and "them": class, mobility and identification in a new town,' *Sociology* 36 (1), pp. 171–93.
- Struthers, T. (2003) 'The redevelopment of Salford Quays, Greater Manchester: its impact on urban and regional development from 1983 to 2003,' *DISP-Online* 154, NSL Network City and Landscape. Zurich: Swiss Federal Institute of Technology. Available at;  
<http://www.nsl.ethz.ch/index.php/en/content/view/full/96/>.
- Thomson, R. and Taylor, R. (2005) 'Between cosmopolitanism and the locals: mobility as a resource in the transition to adulthood,' *Young* 13 (4), pp. 327–42.
- Tyler, I. (2008) "'Chav mum chav scum": class disgust in contemporary Britain,' *Feminist Media Studies* 8 (1), pp. 17–34.
- Urry, J. (2007) *Mobilities*. Cambridge: Polity.
- van Goeverden, C. et al. (2006) 'Subsidies in public transport,' in *European Transport/Transport Europeo* 32, pp. 2–25. Available at:  
[http://www.istiee.org/te/papers/N32/02%20van%20goeverden%20\\_5-25\\_.pdf](http://www.istiee.org/te/papers/N32/02%20van%20goeverden%20_5-25_.pdf).
- Walklate, S. (1998) 'Crime and community: fear or trust?,' *The British Journal of Sociology* 49 (4), pp. 550–69.
- Wrigley, N. et al. (2004) 'The Leeds "food deserts" intervention study: what the focus groups reveal,' *International Journal of Retail & Distribution Management* 32 (2), pp. 123–36.

# Factors hindering the adoption of E-learning in developing countries: Libya a case study - Methodological Perspectives

**Abdulbasit S. Khashkhash, Professor David Eaton**, *School of the Built Environment, University of Salford*

## Abstract

In the last few years, there has been a rapid expansion in Higher Education (HE) in developing countries, especially Africa. As a result, the number of universities in African countries has rapidly increased and these have all been enrolling more students. However, in spite of the importance of E-learning in the advancement and development of HE, E-learning is facing increasing challenges in most universities in these African countries. These challenges include ICT infrastructure, culture, leadership and E-learning strategy, local content, copyright issues, instructors and learners, etc. Moreover, poor economies, gender disparities, etc. are important issues to be addressed in implementing HE. Clearly, without serious consideration and investments in HE and training, the process of developing critical thinking and skills in science and technology could become a mammoth task.

This paper identifies the benefits and challenges facing the implementation of E-learning and outlines advisory measures to overcome the obstacles and barriers. In this manner, governments should define a sound rationale for implementing online learning, recognising that the technology is only a part of the educational transformation process; and that there is also a need to address the lack of infrastructure and trained personnel, as well as the cost of Internet bandwidth and equipment and resistance to change.

## Keywords

E-learning, Developing countries, E-learning Factors, Libya, Research Methodology.

## 1. Introduction

E-learning means the easy use of information and communications technology and the inclusion of the internet in teaching and learning. Yieke (2005) defines E-learning as learning facilitated and supported through the use of information and communications technology. Khan (2000) defines E-learning as encompassing web-based learning (WBL), internet-based training (IBT), advanced distributing learning (ADL), and online learning (OL). With the aid of E-learning, further and higher education can be delivered anywhere and at any time. Colleges and universities are using E-learning to make learning more accessible and efficient. It affords students of all ages and abilities the chance to learn anywhere, at any time and at their own pace. E-learning enhances the learning process by offering a different way of delivering education.

Despite the importance of E-learning in higher education (HE) and human development in Libya, implementation of E-learning is facing a number of challenges in Libyan universities. Those challenges can be summarized as follows: leadership; ICT infrastructure; finance; culture; instructors and learners, and technical expertise. Regardless of these challenges, the future is optimistic, offering innovative and advanced opportunities to implement E-learning. The E-learning program, if established in Libya, will help to solve the problems resulting from a shortage of traditional education institutions that enrol an increasing number of students who wish to study at university, whilst providing them with a chance to learn and promote scientific cooperation and research in order to reach every individual in the community. E-learning would also provide education for those who missed such opportunities earlier, and give them the knowledge of technological developments to continue in other areas of technical development.

This paper includes sections on research aims, objectives, research questions, research methodology, and expected contributions to the body of knowledge, and will benefit the policymakers who are planning and designing the Libyan HE system, as well as the stakeholders; i.e. the teachers and students. It will also help to determine the specification of

main factors in the E-learning process in Libya, and identify the difficulties faced when implementing E-learning and its methods.

## **2. Aim and objectives**

This paper is part of PhD research that aims to develop a framework which identifies the factors affecting implementation of E-learning in HE in Libya; this paper will, however, only focus on the methodology chosen for the research. To achieve this aim the following objectives are laid out: to explore the major barriers that affect implementation of E-learning; to carry out an in-depth investigation to understand the factors affecting adoption of E-learning in HE in Libya; to develop the framework for adoption of E-learning in HE in Libya; to make specific recommendations for overcoming the barriers affecting adoption of E-learning in Libya, and to validate the framework.

## **3. The research questions**

The research questions are:

- Why has there been no implementation of E-learning in HE in Libya until today?
- What are the factors / constraints affecting implementation of E-learning in Libyan HE?

## **4. Critical Success Factors (CSFs) & literature review**

According to Collis and Hussey, (2003) a theoretical framework is a collection of theories and models from the literature, and is a fundamental part of positivistic research as it explains the research questions or hypotheses. However, for a phenomenological study, it may be less important or less clear in its structure. Some researchers have attempted to approach their research where no prior theories exist. In this paper, the factors identified through the literature review and how this would affect implementation of E-learning, create the theoretical framework as follows:

### **4.1 Leadership support**

The transition from traditional delivery methods to the implementation of E-learning environments inevitably involves the management of change (Betts, 1998). The need for support from an organization's leaders in order to begin and maintain any new approach to learning, is addressed in the works of Abdelraheem (2006); McPherson and Nunes (2006). Moreover, Liaw et al. (2007) went further by stating "instructors' leadership is a crucial factor to affect learners' attitudes to implement E-learning". If leadership fails to understand currently emerging futuristic technologies and their potential to develop a vision and strategy to support and enhance learning, acceptance of E-learning will be slow if not impossible (Minton, 2000).

### **4.2 Infrastructure**

Lack of public awareness about ICT and weak data communications infrastructure are two of the main factors affect E-learning (Karmakar and Wahid, 2007). An organization that wants to implement E-learning should attain at least the minimum hardware requirements and the software required. The hardware part of E-learning includes the physical equipment that must be able to supply E-learning, e.g., servers and networks (O'Neill et al., 2003; Ettinger et al., 2006). O'Neill et al., (2003) suggested that the success of technological infrastructure also has implications for the success of virtual learning; malfunctioning hardware or software can both be barriers which can cause frustration and affect the learning process. Valentine (2002) agrees with O'Neill et al. that hardware and tool malfunctions can be greatly detrimental to the effectiveness of E-learning.

Without easy access to appropriate equipment, it is quite hard if not impossible to implement any E-learning. However, Broadbent (2001) states that E-learning does not require a huge infrastructure; even a working internet connection and the provision of enough computers for end-users would be sufficient for an effective E-learning project. James-Gordon et al. (2003) diagnosed that the limitations related to the network capacity of the internet can be a factor when implementing E-learning, and the reason why video and audio transmission can be relatively slow over the internet is that the bandwidth cannot cope with large file sizes. Minton (2000) sees that implementing E-learning requires a minimum technological platform, which includes necessary hardware, adequate telecommunication capabilities and access to software. Universities have to decide what systems, resources and infrastructures would be required to support this type of learning adequately.

### **4.3 Cost**

The consideration of the initial cost as well as the continuing costs of installing, maintaining, using and upgrading technology, and the human capital costs to support E-learning, is very important (Valentine, 2002). Marengo and Marengo (2005) demonstrated that the costs of technological infrastructure include digital content costs, maintenance costs, content hosting costs, hardware and software costs and costs of E-learning staff. Staff costs include tutoring costs, administration and management costs and Expert in Multimedia Technology (ETM) costs. In addition Marengo and Marengo (2005) mentioned that, in cases where the hardware and software supplied by the faculty to circulate the contents are not sufficient, the E-learning evaluation needs to take into account the cost of items, such as the purchase of a server and its relative software. The lack of money can be problematical for implementation of E-learning with continuous labour costs of instructors (Cho and Berge, 2002; James-Gordon et al., 2003; Berge and Muilenburg, 2006).

### **4.4 Resistance to changes**

Resistance to change is one of the important factors for implementation of E-learning (Minton, 2000; Cho and Berge, 2002; Berge and Muilenburg, 2006; Ettinger et al., 2006). This resistance is usual because of the high percentage of illiteracy in some countries (Karmakar and Wahid, 2007). Moreover, Habibu (2003) stated that resistance to change is one of the factors which should be considered when implementing E-learning. It is related particularly to non-technical issues that include academic staff, administrators, and/or managers. It is divided into three main reasons: fear of ICT; lack of time to design, develop and maintain; support for online classes materials, and fear to expose the quality of work. Lecturers are one of the major factors contributing to the success of E-learning. For lecturers, implementation of E-learning programs represents a change in teaching style and materials. The precise nature of the change is difficult to quantify (O'Neill et al., 2003). O'Neill et al., (2003) stressed that human resources should be committed to the project at an early stage and lecturers should be selected based on their attitude towards technology. According to Liaw et al. (2007) the personal attitude is a major factor affecting usage of IT. Understanding the user's attitudes toward E-learning facilities is important for the creation of appropriate E-learning environments for teaching and learning.

### **4.5 Staff education/training**

Besides the lecturer's attitude about using technology, the lack of trained staff to deliver instructions is also important (Valentine, 2002). Liaw et al., (2007) asserted that lecturers should have knowledge and experience of using operating systems, the internet, e-mail, word processing packages, Power-point and basic computer skills to implement E-learning. In this context, experienced and qualified staff are an important factor for implementation of E-learning. According to Levy and Maria (2003) the instructors require training and support to be willing to implement this new teaching paradigm.

#### 4.6 Technical expertise

Lack of personal technological expertise to solve technical problems is one of the main factors for an E-learning program (O'Neill et al., 2003; Berge and Muilenburg, 2006). Valentine (2002) added the one overlooked factor in the success or failure of E-learning programs; i.e. the role that technicians play in E-learning. According to McPherson and Nunes (2006), the need for technical support in using the ICT infrastructure as well as good maintenance of this infrastructure is the main factor in implementing E-learning. It may require a project manager, software programmers, multimedia experts, instructional developers and others, including information systems experts. Even more, Marengo and Marengo (2005) specified that experts in multimedia technology have to support teachers in the activities of organization and management, and to some extent the development of E-learning courses as well. Moreover, Cho and Berge (2002) reported that an organization could not successfully implement E-learning or blend learning without the proper attributes of its infrastructure and making technical expertise and technical support easily available.

The critical success factors from literature which affect implementation of E-learning is described below:

<b>Research [Publication Date] publishing</b>	<b>Factor</b>
Abdelraheem, [2006]. McPherson and Nunes, [2006] Minton, [2000] Liaw et al. [2007] Betts, [1998].	<b>Leadership</b>
McPherson, [2006] O'Neill et al., [2003] Ettinger et al., [2006]. Valentine, [2002] Broadbent, [2001] James-Gordon et al. [2003] Minton, [2000] Karmakar and Wahid, [2007].	<b>Infrastructure:</b> -Hardware -Software -Internet -Network -Support Tools -Server
Valentine, [2002]. Marengo and Marengo, [2005] Cho and Berge, [2002] James-Gordon et al., [2003] Berge and Muilenburg, [2006].	<b>Cost:</b> -E-learning staff -Content Hosting Cost -Digital content
Minton, [2000] Cho and Berge, [2002] Berge and Muilenburg, [2006] Ettinger et al., [2006]. Liaw et al., [2007] Karmakar and Wahid, [2007]. Habibu, [2003] O'Neill et al., [2003].	<b>Resistance to changes:</b> -Academic Staff -Administration

Levy and Maria, [2003] Liaw et al., [2007] Valentine, [2002]	<b>Staff Education/Training</b>
Cho and Berge, [2002] Valentine, [2002] McPherson and Nunes, [2006] Marengo and Marengo, [2005] O'Neill et al., [2003] Berge and Muilenburg, [2006].	<b>Technical Expertise:</b> -EMT -Software Programmer -Information System -Project management

## 6. Methodology

The two main methodological research approaches are deductive and inductive. Saunders et al. (2007) encourage combining deductive and inductive approaches within the same piece of research. The two combined research approaches (inductive and deductive) will be selected for this research to achieve the objectives of the study. In this research a theoretical framework is first derived from the existing literature which will be investigated in the case study organization (deductive approach). Subsequently the author expects to incorporate the findings from case studies into existing theory (inductive approach).

### 6.1 Selection of case study

Rather than a random sampling strategy, a single case study approach will be adopted and a theoretical sampling strategy will be explored and chosen for appropriate selection of the case. In the Libyan Education sector, the Ministry of Higher Education is used as an effective case study which is a unique and important case and meets enough of the criteria for the study. The following are reasons for selection of the case study:

1. The Ministry of Higher Education controls all the universities in Libya. It has all higher education programs and is the only Libyan organization that is specialized and representative in Higher Education;
2. The Ministry of Higher Education has well educated and professional people who have had enough knowledge of technology;
3. Accessibility to the selected case study will be straightforward for the author as the Ministry of Higher Education is situated in Tripoli, Libya's capital city. This allowed the researcher to contact the sector easily, considerably reducing travel time and cost.

### 6.2 Sampling of case study

The case study was conducted in Libya during the period from August to September 2008. All the interviews were carried out in the interviewee's offices, in order to have the opportunity to access the relevant documents. The number of interviewees in the case study organisation was 20; the respondents represent top management, deans, academic staff and support staff. Covering the three managerial levels enhanced the validity through triangulation of the data gathered from different points of view from the respondents. Face-to-face open-ended questions allowed the interviewer to gain in-depth information. The time allocated for each interview fluctuated; in general, the average time was about forty-five minutes to one and a half hours.



## **7. Findings and discussion**

The findings showed that the leadership and professional opinions confirm the advantages of E-learning; however, the changes require strong and supportive leadership as well as changes in the organizational structure and culture. On the other hand, in order to drive the strategy forward within their institutions the educational leaders should be convinced of the benefits of E-learning.

The need for support from HE leaders to obtain the funding for technology, staffing, and other resources, is necessary to start and maintain any new approach to learning. An institution that wants to implement E-learning should fulfil at least the minimum hardware and software requirements in order to operate this hardware. The hardware part of E-learning includes the physical equipment that must be able to supply E-learning. Without easy access to appropriate equipment, it is difficult to implement any E-learning programme.

The HE educational institutions have faced several barriers posed by the rapid evolution of ICT, which form the backbone for E-learning. Universities should consider the costs of installing, maintaining, operating and upgrading technology to support E-learning. In addition, the human capital costs should also be factored in to calculate the true cost of E-learning.

According to interviews conducted by the researcher, resistance to change is one of the factors which should be considered when implementing E-learning, and it usually occurs because of a high percentage of ICT illiteracy among senior staff. In some faculties, resistance to change factors were confirmed and were found to be for the following main reasons: fear of ICT; lack of time to design, develop and maintain; support online class materials, and fear of exposing the quality of staff's work.

Lecturers need training and support to be willing to adopt this new teaching paradigm. The instructors should be trained both to take advantage of their experience, and to be able to adapt that experience to the new environment of E-learning. Therefore all staff involved in E-learning require a wider scope of knowledge and technological skills, such as the ability to use multimedia, email and Internet. Lecturers should be trained to use these and be able to adapt them to the new environment of E-learning.

Finally, the need for expertise in solving technical problems is one of the main factors when introducing an E-learning program, and HE in Libya needs to provide technical support in using the ICT infrastructure, as well as infrastructure maintenance. Universities cannot successfully implement E-learning or blend learning without the availability of technical expertise and adequate infrastructure support.

## **8. Future work**

Supplementary work should be conducted in different educational institutions for generalisation of data, especial in rural universities. The attitudes of students and new staff should be counted towards implementing E-learning.

## **9. Conclusion**

The failure of Libyan HE to accept and develop a vision and strategy to understand the current emerging advanced technologies, will seriously undermine any potential to implement the E-learning Process. E-learning requires a minimum technological platform, which includes necessary hardware, adequate telecommunication capabilities and access to software. The costs of technological infrastructure is to contain digital-content costs, maintenance costs, content hosting costs, hardware and software costs and costs of E-learning staff. Staff costs include tutoring costs, administration and management costs and Experts in Multimedia Technology (ETM) costs. Lecturers should be selected based on their experience, potential and outlook towards new technology. The expert in multimedia technology has to support teachers in the activities of organization and management, and to some extent the development of E-learning courses as well. The implementation of E-learning programs represents a change in teaching style and materials.

This paper has contributed to the discussion of the main success factors and methodology used to implement E-learning.

## References

- Abdelraheem, A., Y., (2006) "The implementation of E-learning in the Arab University: Challenges and opportunities". *In Proceeding of APRU DLI Tokyo, Japan.*
- Abouchedid, K., and Eid, G.M., (2004) "E-learning challenges in the Arab world: revelations from a case study profile" *Journal of Quality Assurance in Education*. Vol12, No.1, PP. 15-27.
- Collis, J., and Hussey, R. (2003), *Business research a practical guide for undergraduate and postgraduate students*. Palgrave Macmillan.
- Khan, B., H., (2000) "A framework for web-based learning". Englewood Cliffs, NJ: Educational Technology Publications.
- Porter, M.E., Yergin, D., (2006) "National Economic Strategy: An Assessment of the Competitiveness of the Libyan Arab Jamahiriya" Cambridge Energy Research Associate (CERA) PP. 1-189.
- Saunders, M., Lewis, P. and Adrian, T. (2007) "Research methods for business students" (4<sup>th</sup> ed.), Prentice Hall, London.
- Twati, J.M., (2006) "The impact of organizational culture innovation on the adoption of IS/IT: the case of Libya". *Journal of Enterprise Information Management*. Vol. 19, No.2, PP. 175-191
- Yieke, F., A., (2005), "Toward Alternatives in Higher Education: The Benefits and Challenges of E-learning in Africa", CODESRIA Bulletin No. 3&4, PP. 73-75
- Yin, R. K. (1994) "Case Study Research Design and Methods" 2<sup>nd</sup> Edition, Sage Publications, Newbury Park.
- Yin, R. K. (2003) "Case study Research: Design and Methods", 3<sup>rd</sup> Edition, Sage publications, Inc.

## Time for Magic: An Investigation into Jamie Reid's Druid Ancestry

**Vicki Maguire**, *Liverpool School of Art and Design, Liverpool John Moores University*

Jamie Reid (b. 1947) is a British artist with connections to the Situationist International. He is best known for his work for the Sex Pistols during the mid- to late 1970s, which featured letters cut from newspaper headlines. This work helped define the image of Punk, particularly in the UK. However, as demonstrated by projects such as the ongoing Eight Fold Year Project, Celtic Surveyor (1989–1992), and his innovative interior design at the Strongroom Recording Studios, London, it is clear that Reid's work both past and present has been influenced by the enduring themes of Druidism, Shamanism, Pagan symbolism, magic, and the forces of nature. These themes are expressed, for example, in Reid's creation of inspirational interior environments based upon Pagan and Druidic symbols, signifying a revolt against the repressive nature of conventional architecture. They are also evident in the earlier graphics for his radical political magazine *Suburban Press*, published during the 1970s, many of which were also later reused for the Sex Pistols. The overall aim of my PhD research is to reassess Reid's body of work from the past thirty years, locating it within the framework of a politicised practice that can be traced back to Reid's family background and attitudes towards politics and religion, with the artist's spiritual concerns being used to explore alternative directions for the future.

The reoccurring themes of Druidism, social and political reform, a strong bond with the natural world, and an interest in alternative lifestyles are echoed explicitly in Reid's ancestral heritage, most notably in the life of his great uncle, the post-Edwardian social reformer George Watson MacGregor Reid (c.1862–1946). Through an exploration of Reid's links with Druidism, it is my intention to place his practice in the wider context of British radical dissent and alternative spirituality, shifting the focus from his role as art director for the Sex Pistols and making him part of an alternative lineage. I shall begin with an in-depth investigation of the life of Reid's great uncle George Watson MacGregor Reid, editor of an alternative health and lifestyle journal and head of the Druid Order in England at the beginning of the twentieth century.

This paper contends that far from standing alone, Reid's exploration of alternative spiritual and political directions for the future and his deeply held respect for the natural world and its innate power have ancestral precedents in a family steeped in the kind of spiritual socialism that flourished as a result of the new freedoms afforded by the end of the Victorian era.

Through a series of original, first-hand interviews with Jamie Reid, I have gained unique access and unparalleled cooperation with the artist and amassed a wealth of new information relating to Reid, his artistic career, and his unconventional family background. During the interview process, the artist discussed in great length the impact of this heritage upon his practice, in particular the influence of Druidism and its relation to the activities of his great uncle George Watson MacGregor Reid, as well as the family's links to socialism and their later involvement in the Campaign for Nuclear Disarmament.

Reid's grandparents appear to be the first relatives that could be said to embody this somewhat unconventional approach to life. Little information exists about earlier generations of the Reid family, which can be attributed, in part, to a lack of written genealogical records. Reid expressed a desire to delve deeper into the family's history "because there are such big gaps." He noted that by the time he and his brother Bruce were born, his grandparents had already passed away (Maguire, 2008).

In Reid and Jon Savage's 1987 collaborative book *Up They Rise*, the artist discusses the fact that his sea-faring paternal grandfather Robert was a gun-runner for the Chinese during the Boxer Rebellion (Reid and Savage, 1987). This uprising of the Chinese secret society I Ho Ch'uan (Righteous Harmonious Fists), known to Westerners as Boxers, lasted from November 1899 to September 1901. It saw the world's eight most significant military powers—Austria-Hungary, France, Germany, Great Britain, Italy, Japan, Russia, and the United States—join forces against this common enemy. The Boxers claimed to be under the protection of ancient Chinese gods in their fight against the influence of foreign "devils." Their followers used incantations and magic rituals performed by shamans to create a "shield" of protection. By the spring of 1901, the Boxers were out of control; missionaries were killed, churches burned, and railroads destroyed as a protest against foreign rule. Numerous Boxers were either executed

or exiled to placate the allies, and peace returned on 7 September of that year (Bodin and Warner, 1979). Reid's grandfather later died in France from wounds received during this time. His young son John Finlay MacGregor-Reid—Reid's father (b. 1908)—was to be cared for by his uncle George, described by Ronald Hutton in his 2007 book *The Druids* as a "tall, massively-built man with a luxuriant moustache" (Hutton, 2007: 172). Reid described his great uncle as an individual "shrouded in some mystery" (Reid and Savage, 1987).

The hopeful fusion of alternative politics and spirituality was a key feature of the Druid order. The belief system of Druidism would later prove attractive to Reid's great uncle George Watson MacGregor Reid, who was the kind of intriguing and dynamic individual that historians and biographers often describe with the clichéd phrase "a colourful character" (Piggott, 1968: 181). Even though sixty-two years have now passed since his death in 1946, much mystery, myth, and rumour still envelope the life of this Chief Druid and social reformer of the post-Edwardian period, a significant amount of which has been generated by the man himself. In his 1968 book *The Druids*, Stuart Piggott lists MacGregor Reid as a friend of the Irish writer George Bernard Shaw, as well as the reputed inspiration behind the Australian novelist Guy Boothby's occultist anti-hero Dr. Nikola, although concrete evidence for these claims appears to be lacking (Piggott, 1968). Even most detailed and intellectually rigorous study of MacGregor Reid's life to date—Dr. Adam Stout's *Universal Majesty, Verity and Love Infinite*—recognises that MacGregor Reid was guilty of "constantly shaking up the kaleidoscope, creating wild and contradictory claims about his past that might have been designed to throw all wannabe biographers off the scent" (Stout, 2005, p. 4).

A large body of Jamie Reid's work draws directly upon this heritage of socialist politics, mysticism, and spiritual engagement with the natural world. Indeed, Reid has described himself as a "socialist Druid" or "lapsed Druid," amongst other things (Reid, 2006). It is therefore important to look beyond the myth and contradictions surrounding George Watson MacGregor Reid in an attempt to decipher and document the activities of this significantly inspirational figure.

Even his date and place of birth is uncertain. According to the Scottish registry, a male by the name of George Watson Reid was born to a nautical family in 1862 in Anderston, Lanarkshire (Stout, 2005: 5). This appears plausible, as the town of Anderston is in close proximity to the docks on the River Clyde, an area that saw the growth of industries such as shipbuilding, engineering, and iron-founding during the nineteenth century (Cooper, 2006). However, the man himself never confirmed this during his lifetime. Instead, he caused confusion by claiming at various stages of his life to have been born on the Isle of Skye or in India, in years ranging from 1850 to 1854/5 (Stout, 2005). MacGregor Reid's childhood and early career remain similarly obscured; Dr. Adam Stout traces Reid's path from fisherman to seaman to docker, although again exact dates and other details have proven elusive for researchers (Stout, 2005). This is problematic from an historiographical point of view, but it also proves useful as it goes some way towards revealing the very nature of MacGregor Reid's personality and his desire to perpetuate a certain image and maintain an air of intrigue—an approach to life which, at its worst, may raise accusations of charlatanism (Hutton, 2007).

In any case, MacGregor Reid first appears in the public consciousness around 1890. His first incarnation is that of a radical political figure, agitator, and trade unionist. He campaigned for dockers' rights at ports in both Britain and America as a delegate of the National Amalgamated Sailors' and Firemen's Union. After spending six months in Hull, MacGregor Reid established branches of the union in both New York and Boston before internal union wrangling brought an abrupt end to his career. He talks about being "badly wounded in battle" sometime around the beginning of the 1890s, which Stout interprets as alluding to a mental scar rather than the physical one. MacGregor Reid later attributes the injury to the time he supposedly spent serving as a "medical officer to Chilean revolutionaries"—perhaps one early example of his tendency to embellish or re-create past events in order to disguise any personal shortcomings or humiliations (Stout, 2005: 8).

Moving on from this experience, MacGregor Reid's commitment to radical politics gained intensity and he would later become one of only a handful of people to have attempted to stand for both the House of Representatives in the U.S. and the House of Commons in Britain (Piggott, 1968). In 1892, he stood for election as candidate for the People's Party (or "Populists") in New York, an organisation that declared war on the rich and wished to place government back in the hands of ordinary American citizens—an obvious choice for a trade unionist (Stout, 2005). However, his bid was unsuccessful, as Populist support was essentially

confined to rural areas. His time in New York was not altogether unproductive, as it is also thought that he became a minister of the Universalist Church during this period. Universalism, formed in the mid-eighteenth century, was an optimistic and tolerant belief system that proclaimed the divine order of the universe, its ultimately beneficent nature, and the triumph of good over evil. Despite having all but died out in Britain by this time, it retained a significant following in New England (Stout, 2005). This shift from radical politics to radical spirituality was a process that, according to Ronald Hutton, was more a change in emphasis rather than an abrupt change in beliefs or principles (Hutton, 2007).

MacGregor Reid appears to have returned to England soon after this election defeat. Evidence of what Dr. Adam Stout called his "small-u universalism" can be found in his Populist/Anarchist pamphlet, *The Natural Basis of Civilisation*, published by the Proletarian Publishing Company of Clerkenwell, London (Stout, 2005: 13). According to Ronald Hutton, this was the first written testament of MacGregor Reid's three enduring beliefs:

in socialism (with sympathy for anarchism), in natural lifestyles and medicine, and in a universal religion of pacifism, equality and mutual love that blended elements of all the world's great faiths. All were aimed against exploitation, privilege, established and inherited authority, and the ability of the rich and powerful to abuse or use the poor and vulnerable (Hutton, 2007: 172–73).

In addition to highlighting the way in which politics, spirituality, and an exploration of alternative lifestyles became very much entwined during the final years of the nineteenth century, this approach echoes the dual esoteric and political nature of Jamie Reid's own practice in the second half of the twentieth century.

After vanishing from any known records between 1893 and 1906, MacGregor Reid emerged back into public consciousness in September 1906 accompanied by a wife and young son, a private fortune, and a desire to pursue an interest in Naturopathy, a method of physical healing involving natural approaches to health and wellbeing (Carr-Gomm, 2006). He appears to have discovered Naturopathy in New York through Emmet and Helen Densmore, a husband-and-wife team of doctors who specialised in treating obesity and who were also authors of an influential 1892 book entitled *How Nature Cures* (Stout, 2005). Naturopathy adopts a progressive attitude to many unconventional issues and beliefs, which remain just as topical and relevant to contemporary society. These include vegetarianism, homeopathy, nutrition, whole foods, outdoor living, and animal rights (Hutton, 2007). These issues were discussed in the flagship journal of MacGregor Reid's British Nature Cure Association, *The Nature Cure*—a monthly magazine with high production values, including a front cover printed in colour. Having returned in 1906 newly married and with newly acquired riches, one might suspect that the cost of such production values was covered by MacGregor Reid's wife. However, Dr. Adam Stout suggests that

like those too-prosperous ex-colonials who turn up in Sherlock Holmes stories, living comfortably but fearfully in Home Counties mansions, MacGregor Reid truly had a shady past, and was now living off gains ill-gotten during those thirteen silent years (Stout, 2005: 18).

Despite the half-humorous tone of this throwaway comment, along with his stated commitment to presenting the "bone-dry facts," it could be said that Stout—perhaps unknowingly—is in fact going some way to perpetuate certain myths surrounding MacGregor Reid's life (Stout, 2005: 4). The fact that Stout has chosen to make such a romantic claim, despite having some evidence of a decidedly more plausible yet mundane explanation, hints at the way in which some biographers still find it difficult to resist what Ira Bruce Nadel describes as "the effect of an arresting image replacing a tedious detail" (Nadel, 1984: 163).

Indeed, MacGregor Reid's very name is thought to have been a partly artificial construct. As previously mentioned, records show that a male by the name George Watson Reid—likely to be the man himself—was born in Lanarkshire in 1862. It was only after his re-emergence in 1906 that George Watson Reid had acquired the extra name of "MacGregor" (Stout, 2005). Research conducted by Stout has shown that the addition of the name MacGregor may have

been a result of his developing interest in the occult, in particular his involvement with one of the many mysterious esoteric societies that started to emerge in Britain towards the end of the nineteenth century. This group of organisations included the oriental mysticism of Madame Blavatsky's Theosophical Society, which was formed in 1875 by the Russian Helena Petrovna Blavatsky. The latter, incidentally, was a close friend of MacGregor Reid's acquaintance Dr. Helen Densmore (Stout, 2005). While critics would later accuse the society of being nothing more than a personality cult, it started out with honourable aims, including the desire to form "a universal brotherhood of humanity" and to "investigate the unexplained laws of nature and the psychical powers latent in man" (Ashe, 2007: 206).

Set up partly as a Western-centred response to this society, The Hermetic Order of the Golden Dawn was a closed group specifically concerned with the study of ritual magic. Ronald Hutton describes it as the "greatest society of ceremonial magicians in nineteenth-century Britain" (Hutton, 2007: 173). The name "MacGregor" had already been adopted by at least four Golden Dawn members, including joint founder and leader Samuel Liddell MacGregor Mathers, praised for his 1887 translation of Knorr von Rosenroth's *The Kaballah Unveiled* (1684) (Soror, S.J., 2000).

Mathers himself was an individual who was not adverse to applying creative license to his own biographical details. Although born in London in humble circumstances, Mathers later claimed to be a descendent of Alasdair MacGregor of Glenstrae, chief of the Clan Gregor. Alasdair led his men to victory in the Battle of Glen Fruin in 1603 and was murdered later that year (Innes, 2007 [1919]).

Described by his close friend J.W. Brodie Innes as a "Celt of the Celts" (*ibid.*), Mathers also took to regularly wearing Highland dress and changed his wife's first name from "Mina" to the more Gaelic "Moina"; Mathers was nothing if not thorough in his embracement of his supposed Celtic heritage (Stout, 2005: 15). Despite speculation as to whether the two men had ever met in person, it seems that MacGregor Reid held Mathers in extremely high regard as a master of cabbalistic teaching, and that he adopted the name "MacGregor" in honour of this spiritual influence (Hutton, 2007). This was occasionally changed into the double-barrelled MacGregor-Reid and—after developing an interest in Naturopathy—it was often brazenly preceded by the word "Doctor" (Stout, 2005).

September 1906 saw the launch of MacGregor Reid's journal *The Nature Cure*, a publication which aimed to appeal to "all Food Reformers, Vegetarians, Temperance Enthusiasts, Anti-Vivisectionists, Anti-Vaccinists and believers in Simple Life ideals" (*Nature Cure Annual*, 1907–8: xiii). It featured a range of articles on contemporary issues concerning alternative lifestyles written by individuals such as Richard Metcalfe, who ran the London Hydropathic Establishment. Each issue also contained one of MacGregor Reid's idiosyncratic, increasingly bizarre editorials, as well as many advertisements for Naturopathic products (Carr-Gomm, 2006). *Nature Cure Annual* was also published during the autumn of 1907.

To start with, *The Nature Cure* showed all the signs of becoming a success; a limited-liability company was set up by MacGregor Reid and his business manager John Shaw. The offices of the magazine soon moved from Burgess Hill to an illustrious Fleet Street address, although it is unclear how this move was funded (Stout, 2005).

Even though *The Nature Cure* appeared to cater to the burgeoning interest in alternative lifestyles, it seems that readership was small and did not live up to MacGregor Reid's grand expectations, which included plans for a horse-drawn caravan that would spread his message through British towns and cities. John Shaw eventually pulled out of the venture; perhaps feeling disillusioned after a number of such appeals for funds produced little response from readers (Stout, 2005).

Within three years, MacGregor Reid's focus—both journalistically and personally—had shifted from healthy living and Naturopathy to healing of a more spiritual nature (Hutton, 2007). *The Nature Cure's* content had become increasingly esoteric, with most of this eccentricity being penned by MacGregor Reid. The British Nature Cure Association had now become The Universal Bond of the Sons of Men, an organisation with bases in both Clapham and Leamington Spa (Hutton, 2007). The Universal Bond reflected MacGregor Reid's personal and ever-changing spiritual interests, which ran from esoteric Christianity and Druidry to Eastern religions such as Buddhism and Zoroastrianism. Members of the Hermetic Order of the Golden Dawn had also supported the cause of the Senussi tribespeople, a group of Muslim mystics

living in the Libyan Desert, who were then under attack as a result of Italian colonial expansion (Carr-Gomm, 2006).

Most of these spiritual interests appear to have been cultivated as a result of the influence of particular individuals such as the “Mahatma” Sri Agamya Guru Paramahansa, a Hindu who visited London from India and extolled the virtues of yoga. However, deeper research indicates that MacGregor Reid may have been somewhat naïve or haphazard in his choice of spiritual inspiration, and he would often quickly discard influences in favour of new projects. This was perhaps a positive approach in the case of Sri Agamya Guru Paramahansa. A *New York Times* article of 12 July 1908, entitled “Mahatma Goes to Prison,” notes that the guru had been sentenced to four months in prison, “charged with attacking girls who had been lured to his place by advertisements for typewriters” (Anon, 1908).

However, as Dr. Adam Stout explains, MacGregor Reid “was not one to sit for long in someone else’s congregation” (Stout, 2005: 21). The Universal Bond or Brotherhood introduced readers of *The Nature Cure* to MacGregor Reid’s “Holy Book of UMWALI,” standing for “Universal Majesty, Verity and Love Infinite” (Stout, 2005: 21). Incidentally this invocation was also used by Jamie Reid as the title of a 1999 painting..

Although Stuart Piggott unequivocally states that MacGregor Reid was “Chief Druid from 1909” (Piggott, 1968: 183), it could be claimed that his conversion to Druidism was more of a gradual, organic process, taking place from around 1908 to 1912 (Stout, 2005). In 1909, *The Nature Cure* journal appears to have folded, along with the British Nature Cure Association. Both closures were the result, suggests Dr. Adam Stout, of MacGregor Reid’s increasing devotion to spiritual matters rather than a lack of funds (Stout, 2005). The offices of the British Nature Cure Association became the home of The Tribune Publishing Company, which appears to have been set up to enable MacGregor Reid to continue his output of idiosyncratic esoterica. This included *The Path That Is Light*—a lengthy tract published in 1910 under MacGregor Reid’s Buddhist *nom de plume* Ayu Subhadra—and a journal entitled *The New Life*. The latter publication entirely replaced the alternative health and lifestyle content of *The Nature Cure* with musings on spirituality and contemporary world politics. According to Dr. Adam Stout, MacGregor Reid’s journal particularly championed the cause of the oppressed “Islamic peoples of the Balkans and North Africa” (Stout, 2005: 26).

Taking the place of the British Nature Cure Association was the South London Temple of the Universal Bond, which by 1913 had taken as its headquarters a large villa just south of Clapham Common (Stout, 2005). This was to become the new centre of MacGregor Reid’s Universalism; at the same time he also became the founding minister of the South London Universalist Church (Druidry.org, 2005). It could be suggested that, after much exploration of world religion and various associations with often dubious spiritual figures, MacGregor Reid had finally reaffirmed his belief in the kind of liberal, universal religion he had discovered earlier in life. As Dr. Adam Stout suggests, this may have been around 1890, when he was organising dockworkers in the U.S. However, Alistair Bate believes that he may have come across Universalism “as a young man living in Glasgow, during the last days of the ministry of the Reverend Caroline Soule, who ministered in both Dundee and Glasgow” (*ibid.*).

MacGregor Reid thus chose to eschew world faiths—both conventional and more leftfield ones—in favour of a belief system that distilled the essence of those faiths into a search for universal love, truth, and meaning. However, one creed managed to retain MacGregor Reid’s attention and favour—namely, Druidism (Stout, 2005). Druidism was believed to have been the ancient British representation or embodiment of the biblical concept of an original true and universal faith, a stance taken up by many nineteenth century writers (Hutton, 2007).

Hutton goes on to explain that “the biblical tradition had been given a recent post-Christian restatement by Sir Norman Lockyer’s assertion that astronomy proved the existence of a uniform archaic faith focused on the sun” (Hutton, 2007: 174). Like many world religions—no least Christianity—MacGregor Reid’s own Universalism equated light with divine knowledge and the search for truth. This was echoed in the Druid celebration of the summer and winter solstices, and the Awen symbol created by the iconic Welsh neo-Druid Iolo Morganwg (1747–1826), the creator of the late-eighteenth-century Druid revival (Icons.org.uk, 2009). The symbol features three rays of light emanating from three separate points surrounded by a circle or series of circles. It is said to represent illumination, inspiration, and wisdom (Carr-Gomm, 2006).

Thus for MacGregor Reid and his followers, there was nowhere more fitting to congregate and celebrate their religious beliefs than Stonehenge, a prehistoric monument in the county of Wiltshire with a growing national profile based upon its believed ancient function as a solar temple (*ibid.*). In the latter decades of the nineteenth century, large and good-natured crowds would gather there annually to watch the midsummer sunrise. As a result of this growing reputation, a site that was once isolated, little-visited, and open to all was transformed from around 1900 onwards into an iconic location actively explored by groups of tourists who were now forced to pay an admission charge to a custodian (Hutton, 2007). This coincided with a change in ownership from private hands to landowner Sir Edward Antrobus. Antrobus's implementation of an admission fee in 1900 may have also been a result of the fall of one of the upright stones as well as the government's refusal to purchase the site (Piggott, 1968).

The neo-Druid association with Stonehenge predated any involvement by MacGregor Reid. The society started off without any overt religious connotations. It is described by Peter Ackroyd as "a convivial society, not a religious assembly" (Ackroyd, 1995: 98). In 1781 a London carpenter and builder by the name of Henry Hurle founded the Ancient Order of Druids, a secret society inspired by the Freemasons, with many privileged and titled members (Piggott, 1968). According to Ackroyd, the order was established in a London ale-house just a few yards from William Blake's house on Poland Street (Ackroyd, 1995). Hurle's Ancient Order of Druids was just one result of the popular and scholarly interest in ancient British history and religion in the eighteenth century—the context in which Blake was operating as an artist and poet. According to Ackroyd, this "new antiquarianism" was

actively rewriting British history—from Cooke's *The Patriarchal and Druidical Religion* to Stukeley's *Stonehenge: A Temple Restored to the British Druids*. Welsh "bards" met on Primrose Hill (there was supposed to be a tribe of Welsh-speaking Indians descended from lost patriarchs) and parodies of "Druid" literature were already appearing (*ibid.*: 98).

In 1833, a breakaway charitable institution, The United Ancient Order of Druids, was formed after internal wrangling over the order's purpose. After this split, the original Ancient Order of Druids continued along the path of mysticism, incidentally accepting a young Winston Churchill as an initiate in August 1908 (Piggott, 1968).

In 1905, seven years before MacGregor Reid and the Universal Bond descended upon the Stones, the Ancient Order of Druids held their first grand ceremony there. This was an initiation of 259 people, with around 700 members present in total. This was carried out in what had now become the full ritual Druid dress of flowing white robes and bushy white false beards (Icons.org.uk, 2009). Members also carried wands and staves, with those of officers being topped with golden sickles. This seems to have been a sedate and well-organised affair, with luncheon being served in a marquee beforehand, accompanied by the strains of a Druid brass band. Tea was served afterwards, with the Druids being treated to "four kinds of the best cake" (Hutton, 2007: 151). The ceremony was conducted with the full support of owner Sir Edmund Antrobus (*ibid.*). Ronald Hutton explains that the order had

carefully applied in advance for permission to use Stonehenge, avoided the turbulent and crowded period of the summer solstice, incorporated and flattered Antrobus in the course of its rites and (to judge from its later recorded practice) made the occasion handsomely worth his while in financial terms (*ibid.*: 175).

This rational approach could not have been more at odds with how MacGregor Reid would conduct proceedings. As mentioned previously, his transformation from Universalist minister to Druid figurehead was a gradual one. In 1912, he took four members of the Universal Bond to the Stones to celebrate the midsummer solstice under his pseudonym Ayu Subhadra. He was robed in a distinctly Eastern style and wore a white turban, as shown in a series of photographs from the time. This unconventional-looking group held two services—which at this time still contained references to Zoroastrianism and Islam—and gave themselves Persian, Turkish, and Sanskrit titles (*ibid.*). True to form, MacGregor Reid had not warned Antrobus



about his arrival but instead attempted to win favour with the journalists present, relishing the publicity his group were generating and using the opportunity to hand out copies of the Universal Bond's declaration (*ibid.*).

MacGregor Reid returned with nine followers in 1913, this time in the guise of another spiritual hybrid, the Dastur Tuatha de Dinaan (Stout, 2005). Despite being informed that Sir Edmund Antrobus would no longer tolerate any political or religious meetings at Stonehenge, MacGregor Reid managed to gain admission and hold his service, albeit after a fracas with both caretakers and police. After a successful and inspiring meeting, he promptly brought down the first of many thunderous curses on the head of Antrobus, proclaiming, "In grief and sorrow I call down the curse of Almighty God, and of the tyrannical. . . . Response is sure" (*ibid.*:32). Similar scenes occurred in both 1914 and 1915, when MacGregor Reid demanded his money back from Antrobus and was forcibly removed from the site by police. According to Ronald Hutton, these episodes were well documented by the newspapers of the time, although unfortunately for MacGregor most reports were unsympathetic, with one journalist labelling the group a "brawling party of cranks" (*ibid.*: 34). Despite this, when Antrobus died of natural causes soon afterwards at the age of 66, MacGregor Reid and his followers claimed victory and declared their curses a success (Icons.org.uk, 2009).

It was around this time that MacGregor Reid began to identify directly with Druidism, calling himself "The Last of the Druids" (Stout, 2005: 32). However, it is evident from research that there appears to be no clear point at which the Universal Bond mutated into the Ancient Order of Druids. The process seems to have been an evolution taking place over a number of years, and definitive documentation of this process is lacking. Philip Carr-Gomm states that this did not take place until the 1940s or 1950s (Carr-Gomm, 2006). Dr. Adam Stout goes as far as to say that "there was little love lost between the two orders," and that the Ancient Order of Druids "took great pains to distance itself from the Universal Bond when the latter's activities generated bad publicity" (Stout, 2008: 144).

As Ronald Hutton explains, "The true story of the development of the Universal Bond between 1906 and 1912 was obliterated—at least to outsiders—to be replaced by a claim that it was the heir to a continuously existing Druid tradition that had come down from ancient times" (Hutton, 2007: 180). The Universal Bond now claimed to have been in existence since the seventeenth century. From a period of around ten years, from the mid-1910s to the mid-1920s, the number of members increased to around fifty. Their eastern dress was replaced by more conventional white robes, and members were now organised into lodges similar to those of the Druids and Freemasons (Hutton, 2007). In 1918, MacGregor Reid told a large congregation that "Faith in the Druid God will . . . make wars cease" (Stout, 2005: 34). Around this time MacGregor Reid's followers also began to call themselves "An Druidh Uileach Braithreachas," translated as the Universal Druid Brotherhood (Stout, 2005). It was obvious that despite their differences, Universal Bond and the Ancient Order of Druids were becoming increasingly aligned in both ideology and appearance.

Between 1916 and 1918 the Bond held a number of peaceful ceremonies aided by the kindly nature of local businessman Cecil Chubb, Stonehenge's new owner. However, later on in 1918 Chubb donated Stonehenge to the nation and it was brought under control of the government's Office of Works, which retained both the admission fee and the problematic former caretaker Sir Edmund Antrobus (Stout, 2008). While the government's Inspector of Ancient Monuments allowed the Universal Bond to continue to worship at Stonehenge, the group were not exempt from payment of the admission charge or the wrath of the aforementioned caretaker and others who shared his stance. The years that followed would be a difficult and dramatic period for both MacGregor Reid, who at times proved to be somewhat of a thorn in the side of Lloyd George's government. Similar clashes followed from 1924 onwards with the next Prime Minister Ramsey McDonald, this time in relation to recent archaeological explorations of the site (Stout, 2008).

During the years of the First World War, MacGregor Reid had also returned to his socialist roots, becoming a pioneer Labour Party supporter and leader of his local Clapham division (Hutton, 2007). His speeches were aimed firmly at working people, who welcomed his denouncement of the rich and powerful in the era of the Great Depression and the General Strike. Logically, he believed that this would win the Universal Bond favour with the new Labour government. However, their hopes were shattered when in 1925 there was another admission fee dispute. MacGregor Reid, now accompanied by his son Robert, incited the

3,000-strong crowd to charge the entrance gate, which yet again led to the involvement of the police (Stout, 2008).

A change in custodian led to another peaceful interlude which lasted until 1932, when the Office of Works denied the Universal Bond the right to distribute its journal—now entitled *The New Life and Druid Journal*—as a response to the official Stonehenge guidebook, written by a local museum-keeper who was keen to disprove the connection between the Druids and Stonehenge (Hutton, 2007). Many of the power struggles and arguments occurring during this period have been documented in detail by Ronald Hutton, and in particular Dr. Adam Stout in his 2008 book *Creating Prehistory: Druids, Ley Hunters and Archaeologists and Pre-War Britain* (Stout, 2008).

As the Universal Bond was unable to defend itself against official propaganda, it seems that the indomitable MacGregor Reid finally admitted defeat or lost interest in his role as Chief Druid. In any case, the Universal Bond never returned to Stonehenge after 1932 and left the dwindling members of the original Ancient Druid Order and other similar groups as the only regular worshippers—a surprisingly uneventful and uncharacteristic end to MacGregor Reid's Druidic associations (Hutton, 2007).

According to Ronald Hutton, by the mid-1930s MacGregor Reid had ceased to speak of Druidism. His beliefs were becoming increasingly conventional, and he gradually realigned his religious attitudes with Christianity (Hutton, 2007). Macgregor Reid's organisation quietly dropped all references to Druidism, becoming "The Universalist Church" and working to forge links with like-minded American churches.

Paradoxically, the later years of MacGregor Reid's life were taken over with one of his earliest preoccupations—Naturopathy. Both the British Nature Cure Association and *The Nature Cure* journal enjoyed short-lived revivals, taking their cue from Stanley Leif's highly successful magazine *Health for All*, launched in 1927. Lief also created what was believed to have been the first health farm at Champneys, a grand stately home near Tring, Hertfordshire (Stout, 2005). This appears to have been one of the inspirations for MacGregor Reid's Royhill Nature Camp at Blackboys in Lewes, Sussex, which functioned from sometime in the 1930s until 1957. By this time MacGregor Reid had married his second wife Alice Biffin, who is also believed to have funded the venture (Stout, 2005).

Royhill consisted of communal halls, dormitories, kitchens, and gardens that attempted some form of self-sufficiency. The camp was described as a place where "men and women in the vanguard movement of politics and religion might come for rest and recuperation" (Stout, 2005: 50). This was achieved partly through a series of baths and other water therapies. Recent research conducted by the Druid Philip Carr-Gomm has also hinted at a Naturist element: elderly members of the Blackboys local history society claimed to have spotted "holidaymakers in their birthday suits" before the war (Carr-Gomm, 2008). However, the unearthing of original Royhill promotional material by Dr. Adam Stout shows that there was an official requirement for bathing suits. At times, MacGregor Reid's principles of healthy living also appear to have been sidelined in pursuit of profit, with meat being served at most meals and cigarettes being stocked in the canteen (Stout, 2005).

MacGregor Reid and his wife moved to Royhill permanently in 1939 and continued to live there despite the British Army requisitioning it for military use later that year (Stout, 2005). He resided at Royhill until his death in 1946.

George Watson MacGregor Reid left a rich and varied legacy. He handed over his mantle of the ministry of the Universalist Church to the politician and journalist Arthur Peacock, and in the 1950s the church became part of the Unitarians. Others who wished to keep the Druid links alive gathered around a Londoner named George W. Smith, whom MacGregor Reid had previously given his blessing (Hutton, 2007). Unfortunately, MacGregor Reid's son Robert believed himself to be his father's rightful successor, and also claimed the name of the Ancient Druid Order for his group of followers; this claim was eventually approved by the Office of Works (Stout, 2005).

After Smith's death in 1954, Robert's order continued to celebrate the solstice at Stonehenge, aided by his diplomatic and genial nature. By the time of his sudden death from a heart attack in 1964, Robert MacGregor-Reid (note the now hyphenated surname) had done much to establish Druid worship of the summer solstice at Stonehenge as a traditional and part of British heritage (Hutton, 2007).

Freud once stated that “Anyone who writes a biography is committed to lies, concealments, hypocrisy, flattery and even to hiding his own lack of understanding” (Nadel, 1984: 178). It would not be outlandish to suggest that similar accusations may also be applied to George Watson MacGregor Reid—a factor that has made my endeavour to dissect the life of this charismatic and eccentric individual even more difficult. In any case, it seems that the creation and perpetuation of myth and rumour has been one enduring theme of MacGregor Reid’s life and work; to acknowledge this fact both before and during this research was the most important task.

Both Ronald Hutton and Adam Stout have discussed the contradictory nature of MacGregor Reid’s character. While both writers have a personal interest in the subject of Druidism, they have also been partial enough to admit that many of MacGregor Reid’s most positive characteristics were often counterbalanced by less appealing personality traits. He believed strongly in peace between nations, yet at the same time did little to reconcile differences closer to home—in particular the thorny and antagonistic relationship between the Universal Order and Stonehenge owner Sir Edmund Antrobus. Many individuals, including Arthur Peacock and George W. Smith, would become loyal and devoted followers. Yet according to Adam Stout, MacGregor Reid was “not at all a team player” (Stout, 2005: 60). Ronald Hutton goes so far as to label him a “bully” (Hutton, 2007: 181).

Despite a tendency to construct fantasies and alternative mystical personas that often led to accusations of charlatanry, MacGregor Reid also seems to have been an individual with generous and charitable ideals. He promoted a doctrine of pacifism, equality, and love, and did much to highlight the unity between world religions, both Eastern and Western. Significantly, the Universal Bond may also be regarded as the group responsible for placing Druidry at the heart of the Stonehenge summer solstice (Stout, 2005).

MacGregor Reid campaigned for social justice both as an individual and as a member of the Clapham Labour Party. He championed an alternative lifestyle that promoted vegetarianism, homeopathic medicine, and a simple, natural approach to all aspects of living (often described as “naturalist” or “simplicitarian”) at a time when such concerns were still regarded as radical and often ridiculed (Stout, 2005).

Ever resistant to the expectations and conventions of mainstream society, MacGregor Reid may be regarded as a countercultural icon. His refusal to pay an admission fee at Stonehenge, and his subsequent removal from the site by police could be said to echo the notorious event of Jubilee night—7 June 1977—when his nephew Jamie Reid, along with Sex Pistols manager Malcolm McLaren and others, were arrested as they accompanied the band on a trip down the Thames aboard the aptly named Queen Elizabeth. Both George Watson MacGregor Reid and his nephew Jamie Reid are individuals who can be defined by their desire to demonstrate new ways of organising our spiritual and political resources. As John Marchant of Isis Gallery, London, explains, “It is this dialectic between gnosticism and dissent that lies at the heart of Reid’s practice and makes him one of the great English iconoclastic artists” (Marchant, 2007). Reid’s artwork addresses and responds to contemporary political and social issues with the same intensity, creativity, and passion that his great uncle George Watson MacGregor Reid conveyed to his followers as a social reformer and leader of the Druids so many decades earlier. This vital relationship offers us a new insight into the life and work of the artist.

## References

- Ackroyd, P. (1995) *Blake*. London Vintage.
- Anonymous (1908) 'Mahatma Goes to Prison' in *The New York Times*, July 12<sup>th</sup>, 1908.
- Ashe, G. (2007) *The Offbeat Radicals: The British Tradition of Alternative Dissent*, London: Methuen.
- Bodin, L.E. and Warner, C. (1979) *The Boxer Rebellion*. London Osprey.
- Carr-Gomm, P. (2006) *What Do Druids Believe?* London Granta Books.
- Carr-Gomm, P. (2006) Weblog. Available at:  
<http://philipcarrgomm.wordpress.com/2008/07/02/detectives-at-work/>.
- Cooper, J.N. (2006) "Anderston Then and Now: A Concise History of a Glasgow District." Available at: <http://www.anderston.ukpals.com/partone17-20.htm>.
- Druidry.org (2005) Available at:  
<http://www.druidry.org/modules.php?op=modload&name=PostWrap&file=index&page=http://www.druidry.org/oldindex.shtml>.
- Hutton, R. (2007) *The Druids: A History* London: Hambledon Continuum.
- Icons.org.uk (2009) Article on Stonehenge. Available at:  
<http://www.icons.org.uk/theicons/collection/stonehenge/features/appropriating>.
- Innes, J.W.B. (2007 [1919]) "MacGregor Mathers: Some Personal Reminiscences." Available at:  
[http://www.esotericgoldendawn.com/tradition\\_macgregorreminiscences.htm](http://www.esotericgoldendawn.com/tradition_macgregorreminiscences.htm).
- Nadal, I.B. (1984) *Biography: Fiction, Fact and Form*. London: Macmillan Press.
- MacGregor Reid, George Watson (ed) (1908) *The Nature Cure Annual and Health & Pleasure Guide*. London.
- Marchant, J. (2007) Commentary on Jamie Reid.  
Available at: [http://www.isisgallery.org/artists/jamie\\_reid.html](http://www.isisgallery.org/artists/jamie_reid.html).
- Piggott, S. (1968) *The Druids*. London: Thames & Hudson.
- Reid, J. (2006) "A Conversation with Jamie Reid," *The Idler* 38. Available at:  
[http://www.isisgallery.org/further\\_reading/jamie\\_reid\\_the\\_idler\\_article.html](http://www.isisgallery.org/further_reading/jamie_reid_the_idler_article.html).
- Reid, J. and Savage, J. (1987) *Up They Rise: The Incomplete Works of Jamie Reid*, London: Faber and Faber.
- Soror, S.J. (2000) "The Hermetic Order of the Golden Dawn in Britain: Samuel Liddell MacGregor Mathers." Available at: <http://www.hogd.co.uk/mathers.htm>.
- Stout, A. (2005) *Universal Majesty, Verity and Love Infinite: A Life of George Watson Macgregor Reid*. East Sussex: The Order of Bards Ovates and Druids.
- Stout, A. (2008) *Creating Prehistory: Druids, Ley Hunters and Archaeologists and Pre-War Britain*. Oxford: Wiley-Blackwell.

# Defining Sustainability in Terms of Independent Filmmaking

**Nico Meissner**, *School of Music, Media and Performance, University of Salford*

This paper is part of my PhD research project, which looks at the internet and the possible models it offers to sustain independent filmmaking. This research project is motivated by the fact that independent film plays vital roles in supporting the commercial film sector and in presenting both alternative voices and rarely told stories.<sup>1</sup> However, it can hardly fulfil those roles in a UK media landscape in which 70% of cinema screens are controlled by only three exhibitors (Macnab, 2009) and five broadcasters share 80% of the television audience (Broadcasters' Audience Research Board, 2009). The internet promises a less gatekeeper-orientated media landscape, in which free access to the medium and its users might empower the independent filmmaker and, as a result, enable independent film to better fulfil the above mentioned roles. But does the internet also offer models to sustain independent filmmaking?

Before looking at how the internet could help sustain independent filmmaking, it is necessary to define what is meant by "sustainability" in this context. However, the existing literature does not provide us with a satisfactory definition, and there is in fact much disagreement over the meaning of the term. One notable debate revolves around filmmaking and its impact on our natural environment. Emma Gardner (2007), for example, attempts to embed filmmaking into the wider issue of sustainable development. Her paper gives advice on how filmmakers can help sustain the environment we are living in.

When debating how to sustain filmmaking itself, the discussion quickly focuses on economic issues. In his presentation "Towards a Sustainable UK Film Industry," Alan Parker (2002) argues that a film simply has to earn back the money for its successor in order to create a sustainable industry. This view is not surprising, since Parker is a full-time filmmaker who sees filmmaking as an income-generating occupation. In contrast, a film scholar like Thomas Elsaesser (2005) uses the term "sustainable filmmaking" rather vaguely and possibly as an alternative to "maintaining filmmaking." Yet even he finds it hard to not focus on economic factors:

But since no West European country can sustain its filmmaking activities without the various subsidy systems put in place during the 1970s and 1980s" (Elsaesser, 2005: 71).

[There is] no longer Adorno's Culture Industry, but the "culture industries" dedicated to generating diverse forms of consumption (different material and immaterial aggregate states of the "work" or "text": videos, CDs, T-Shirts, badges, toys), in order to sustain production (Elsaesser, 2005: 282).

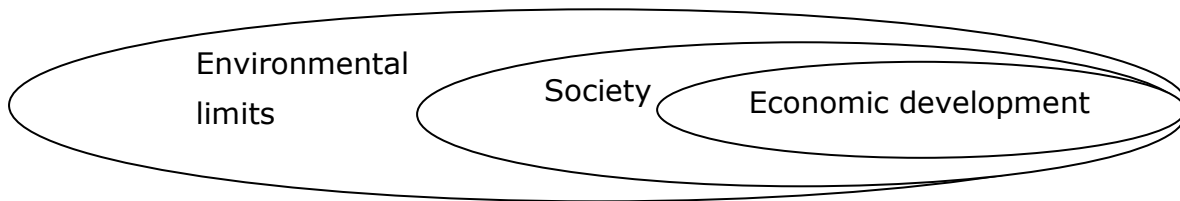
Finally, the popular discussion is also revolving around money as the only means to sustain independent filmmaking (see, for example, Kaowthumrong 2009; Institute for International Film Financing, 2009; Independent Cinema Office, 2001; and ProjectX, 2009).

## **Sustainable development and sustainable filmmaking**

The term sustainability is used in numerous disciplines. It is said to have originated in forestry, where it emphasises survival, as opposed to conditions that are static or even stagnant (Cramer, 2008). Its most prominent usage, however, is within the debate about sustainable development. Here the general consensus seems to be that all economic activities should lead to social progress, but that this must be achieved within the limits set by the environment. This is shown in the Russian Doll model (O'Riordan, 1998):

---

<sup>1</sup> I discuss the importance of independent film in more detail in my PhD dissertation.



Filmmaking is part of our world and as such it has to help sustain the environment in which we live, develop our societies, and maintain economic prosperity. But just what sustains filmmaking? When asking this question, distinguishing between environment, society, and economy does not make much sense; instead, in the context of filmmaking, it is better to replace the concept of “development” within the idea of “sustainable development.”<sup>2</sup>

While it may not be possible to directly transpose the notion of “sustainable development” into “sustainable filmmaking,” it still is very helpful to do so. Sustainable development basically means to ensure growth that does not destroy what it is meant to support. If we now accept that “growth” stands for “development,” “not destroying but supporting” connotes “sustainability.” The purest form of supporting without destroying is to keep something alive. This meaning of sustainability—to keep alive—corresponds with the actual meaning of the verb “to maintain,” as well as with its original meaning in forestry.<sup>3</sup> The question that now arises is simply, How to keep independent filmmaking alive?

The obvious answer, according to both Parker (2002) and a number of the UK Film Council’s initiatives, is that a film has to earn enough money to finance the next film. I want to take a step back and offer an alternative answer. In my view, independent filmmaking will stay alive as long as independent filmmakers make independent films. This includes new independent filmmakers starting their careers and existing independent filmmakers continuing theirs.

This statement is of central importance in my quest for a definition of “sustaining” independent filmmaking. It thus deserves a little explaining. I propose to look at the individual filmmaker rather than at the film sector as a whole. The film sector is made up of a number of individuals, with the filmmaker as its core: without filmmakers, there would be no films, and without films there would be no film sector. Here my approach differs from that of Parker, who sees the financial profitability of a film as a means towards a sustainable industry. I, on the other hand, argue that the individual filmmaker should be the centre of our concerns. Sustaining her leads to her making films, which consequently leads to the survival of the independent film sector.

In order to define sustainable filmmaking, I suggest asking: “How do people become independent filmmakers?” (*career paths*) and, “What do people need to make independent films?” (*needs*). I further assume that a certain kind of motivation is needed for the pursuit of any undertaking. It is hence necessary to also ask, as a question to that of needs: “What motivates people to make independent films?” (*motivation as part of needs*).

The one-dimensional, purely economic notion of sustainable filmmaking found in Parker and in the popular discussion is more concerned with the commercial sector than the independent one. My concern is with the independent sector, which includes a vibrant and flourishing online filmmaking culture. On portals such as YouTube, Vimeo, Jaman, one can find people spending £3,000 of their own money to make a prequel to the *Lord of the Rings* saga, for example (Bouchard, 2009). Indeed, in their spare time many people have made independent films and

<sup>2</sup> In other words, I am more interested in sustaining filmmaking than sustaining economic, social, or environmental development. This is not to say that the latter can be neglected. Ultimately, sustaining our environment and developing societies and economies should be a global aim. There is no doubt that filmmakers, like all artists, can help to do so—by adding to the economy, advertising social development, filling gaps in the knowledge of society, questioning social norms, and campaigning for environmental friendly behaviour. However, in this study, I am more concerned with filmmaking as a microorganism, so to speak, and less with its impact on a macro—i.e., global—dimensions. The latter has been the subject of previous studies (e.g., Gardner, 2007) and needs to be the focus of future investigations. I maintain that independent filmmaking has to be “healthy” itself before it can help “heal” the world.

<sup>3</sup> I want to note that this analogy has been criticised as a truism. However, is not every logical deduction a truism? Moreover, the constant “wrong”/limited use of the term sustainability within filmmaking seems to make it necessary to bring the discussion back to the word’s original meaning.

feature films using stock footage found on the Internet Archive (Kluge, 2008). One can even find filmmakers who have given their documentaries away for free (Geylor, 2009). An economic definition does not offer a direct explanation for the creation of these works. I hence hope to be able to develop an alternative definition that helps us to understand these phenomena. Taking a closer and more nuanced look at the question, "What keeps independent filmmaking alive?" will hopefully lead to a richer understanding of the career paths and needs behind filmmaking.

Since the field of film studies limits itself almost entirely to studies of the film form, I will have to resort to other academic disciplines in order to review literature on the above questions. In the following section, I will examine theoretical perspectives and research findings from sociology, philosophy, and psychology. These fields do not deal with independent filmmaking directly, but rather with arts in general.

## **Literature in related disciplines**

### *The sociological perspective*

The sociological perspective is, unsurprisingly, one that revolves around people as part of social groups. Howard Becker (1982) explains that the production of art is a collective undertaking. In the case of filmmaking, for example, equipment needs to be invented, manufactured, shipped, sold, and purchased. People have to be trained to use the equipment, a film idea has to be developed, and a cast and crew hired. The film needs to be shot, edited, and colour corrected. The sound needs to be recorded, edited, and corrected, and the final cut of the film needs to be printed, copied, and sold. Later the film needs to be exhibited and an audience needs to watch it. The exhibition process is arguably unnecessary,<sup>4</sup> but it is still apparent that a lot of people are involved in the production of art in general and film in particular. If we just look at the end credits of an average film in the cinema, we can easily understand that film production is a collective undertaking, and that the artist is at the centre of this process. She "works in the center of a network of cooperating people, all of whose work is essential to the final outcome" (Becker, 1982: 25).

Becker does not discuss motivations for making art, but he provides us with an idea of what else is needed. According to Becker, making a work of art requires that people:

- cooperate,
- have the necessary talents and abilities, and
- have knowledge of the conventions and restrictions in their artistic sector.

### *The philosophical perspective*

Becker's study explores the "what" rather than the "why" of art production. In *The Act of Creation*, Arthur Koestler (1964) explores the question of why people create things. He concludes that the answer lies in our curiosity and desire to solve problems and explore new things. Koestler goes as far as to say that discovery is more satisfactory than motivations like greed, power, or fame. All of these might be reasons for starting an exploration, but discovery itself leads to bigger satisfaction than money, status, or acknowledgment. Koestler also states that creation is never possible without prior knowledge, and that creation is in fact never innocent but always a result of combining existing knowledge in a new way.

According to Koestler, art thus needs knowledge, i.e., education in a particular artistic sector, in order to get made. Moreover, art is motivated by a will to explore and discover on the part of the creator, someone who is intrinsically motivated by the artwork itself.

---

<sup>4</sup> It might be questionable if art is art before anybody appreciates it. However, a discussion of when art is art and what art is in general does not belong here.

### *The psychological perspective*

Psychologists are, naturally, rather more concerned with job motivation than with the question of what people need to do a certain job. Herzberg et al. (1959), for example, found in their study of factory workers and engineers that five factors mainly explain job satisfaction: achievement, recognition, the work itself, responsibility, and advancement (in this order). Salary is less an explanatory factor for job satisfaction than it is for dissatisfaction with a certain work. Herzberg et al. reason that job satisfaction can be explained by intrinsic rather than extrinsic factors. This means that "getting a job done" (achievement), being praised for one's achievements (recognition), being satisfied with the work itself, being responsible for one's work rather than simply following instructions, and being able to personally develop and advance are all more satisfactory than extrinsic factors such as salary, status, relations to colleagues or superiors, job security or working conditions. Herzberg et al. (1959: 70) concludes that "a sense of personal growth and of self-actualization is the key to an understanding of positive feelings about the job." Centers and Bugental (1966) found that intrinsic motivations are more important in higher occupational levels, whereas extrinsic motivations (such as pay and job security) are more important in lower occupational levels.<sup>5</sup>

Teresa Amabile (1996) looked at creativity in general and came to the same conclusion as Herzberg et al. She found that creativity is motivated intrinsically rather than extrinsically. Extrinsic motivations, such as monetary rewards or recognition, can even hinder creativity because they pressurise the creator. She states that intrinsic motivation maximises creative performance:

The literature review leaves me with the notion that artistic creation first and foremost needs

- knowledge of conventions and restrictions in one's art sector,
- talent,
- ability to make art and
- people to cooperate

and is motivated by

- an internal desire of the artist to explore, discover and create and
- the possibilities for achievement, recognition, responsibility and advancement.

In addition, we have Parker's claim that a film must earn enough money to finance its successor.

Those answers can only be a first hint. Their concern with creativity in general, rather than with filmmaking in particular, helps me to focus my thoughts but does not provide a definite answer to my questions of interest—namely, *How* do people become independent filmmakers? *What* do people need to make independent films? And, secondary to the latter, *what* motivates people to make independent films?

### **Developing a methodology**

The problem I am concerned with differs from the answers provided in the literature in the following respects:

---

<sup>5</sup> When we accept that higher occupational levels are meant to have a higher level of responsibility for the outcome of a task and more managing and supervising duties, the work of a filmmaker can be considered to be at a higher occupational level.



- I am only interested in filmmakers. Filmmakers might have very specific motivations or needs that cannot be generalised for the whole population of artists. This restriction to one form of artistic expression – filmmaking – might enable me to discover much more precise needs for this specific population.
- Besides identifying factors necessary for the creation of independent films, I am also interested in the careers of independent filmmakers. Career paths were not discussed in the literature.

I thus decided to directly confront independent filmmakers and interview them about their career paths and their needs (and motivations) regarding their filmmaking. Answering those questions will provide me with a theory about filmmakers' career paths and needs. This theory will then allow me to define the term sustainable filmmaking.

The data collection process for this part of my research will follow McCracken's (1988) method of the Long Interview, while data analysis will utilise the method of Analytic Induction. I shall use the following pages to describe my methodological approach to data collection and analysis. Before doing so, I feel it is necessary to provide a few words about my approach to the concept of theory and my research philosophy.

### **Data collection**

Ethnographic researches, in line with the philosophy of interactionism, contend that if we try to understand a phenomenon, we should stop looking at it from the outside. We have to dive into a subculture in order to understand it. Becker (1953) studied marijuana users. He wanted to explain people's use of marijuana and did so by finding out how they developed this habit. I will follow this approach and interview filmmakers in order to find an answer to above questions of interest.

In his methodological thinking, Glaser (1992) forbids all literature review prior to data collection, arguing that this will lead to preconceived ideas, which will hinder the researcher's ability to investigate independently. He states that "the dictum in grounded theory research is: There is a need not to review any of the literature in the substantive area under study" (Glaser, 1992: 31). Knowledge in literature, however, is only one source of ideas. There is also the experience and knowledge that the researcher brings to the investigation prior to data collection and analysis. Since I do not see any difference between my own knowledge and the knowledge in existing literature, I decided to break with Glaser's dictum. For McCracken (1988), the literature review is in fact essential prior to the research, as it allows one to see where one's research fits in and to put one's personal preconceptions into perspective.

In addition to the findings of the literature review (see above), I have my own pre-existing ideas of the needs of independent filmmakers. I think, an independent filmmaker needs

- peers who work with, help, influence, criticise, lecture, train, and question the filmmaker, as well as introduce her to new developments, aesthetics, audiences, or peers
- equipment
- money
- time
- ideas
- motivations, which may include:
  - film as an outlet for artistic or political expression
  - the filmmaking process as a form of exploration
  - getting to know or keeping in touch with people
  - developing or maintaining a career in the media
  - earning money with films
  - fame, peer, or audience recognition

It is essential that those ideas, gained from literature review and personal knowledge, do not hijack the data collection or analysis but accompany the process. My research shall not be used to verify or falsify pre-existing concepts but to find out what filmmakers think themselves. Openness to the emergence of new concepts is thus indispensable.

After reviewing literature and one's own ideas, McCracken (1988) suggests developing the questionnaire for the interview. As the research is explorative, the questionnaire will not be set in stone or consist of very detailed questions. It rather lists the topics the researcher wants to cover and offers what McCracken calls Grand Tour questions. Those questions introduce a topic. As all interview questions in this approach, they are non-directive—i.e., they do not suggest any answers. A classical Grand Tour question would be "Tell me about your childhood." According to McCracken, Grand-Tour questions are designed to make the interviewee speak. The interviewer should only interrupt in order to ask for clarification. Apart from this, the interviewee determines where the interview goes. The interviewer listens and makes sure that all topics/areas of interest are covered. If the Grand-Tour questions do not lead to the interviewee covering all topics the researcher was interested in, the researcher will have to use Planned Prompts before going to the next Grand-Tour question. I will use Planned Prompts and the following Grand Tour questions:<sup>6</sup>

For the question, *How do people become independent filmmakers? (career paths)*:

1. Tell me about your filmmaking career; from your first interest in film to your current situation.
  - a. Developing an interest in filmmaking
  - b. Deciding to make films
  - c. Learning about the creative and the mechanical side of filmmaking
  - d. Getting the first film made
  - e. Current situation
  
2. Looking back at your career so far, what do you think were the most important points that brought you to where you are today?

For the questions, "*What do people need to make independent films?*" (*needs*), and "*What motivates people to make independent films?*" (*motivation as part of needs*):

3. How does your typical working day look like (during the periods when you are making films and also when you are not making films)?
  - a. What do you need to make a film?
  - b. Are those things available to you?
  - c. What would you ideally have when you make a film?
  - d. What are your biggest problems when making a film?
  - e. Do you need people, knowledge, skill, money?
  
4. How do you finance your films? And how much money do you earn with a film?
  - a. Do your films earn you a living?
  - b. Do they make back their production money?
  - c. Do they earn the production costs of your next project?
  - d. Does the budget for your next project depend on the financial outcome of the previous project?
  - e. How do you finance your life and your filmmaking?
  
5. What does filmmaking mean to you? Why do you make films?

---

<sup>6</sup> Grand Tour questions are preceded by numbers; Planned Prompts by letters.

- a. What are you trying to achieve by making independent films? When do you know you achieved those things?
- b. Do you want your work to be recognized by others? What form of recognition is important for you?
- c. What is it that fascinates you about filmmaking?
- d. Do you see independent filmmaking as an ends or a means to your career? What are your career aspirations?

6. What situation/event would make you stop making films?

I chose those questions because a) they cover all topics suggested in the literature review and b) they leave enough room to discover new topics, by talking about filmmaking in a general way. Nevertheless, the question catalogue deserves additional explanations.

The block of questions concerned with the career paths of independent filmmakers follows the assumption that "patterns of behavior develop in orderly sequence" (Becker, 1973: 23). I speculate that no one can become an independent filmmaker without developing an interest in filmmaking, wanting to make films, learning about filmmaking, and finally making films.<sup>7</sup>

Those steps are covered by the Planned Prompts of question 1. Question 2 is aimed at allowing the interviewee to summarise the answers to question 1.

The second block of questions covers three topics. Question 3 looks at filmmakers' needs. The Planned Prompts for the Grand Tour question are mainly aimed at covering the findings of Becker (1982) and Koestler (1964) and illuminating the subject from different angles in order to identify needs that were not covered by Becker or Koestler. Question 4 discusses the monetary aspect.<sup>8</sup> Question 5 is concerned with the motivations of independent filmmakers. The Planned Prompts to question 5 follow Herzberg et al.'s (1959) findings for job motivations.<sup>9</sup>

Question 6 plays a slightly different role than the preceding questions. I suspect that it is difficult for artists to talk about their motivations, the reasons behind their work, and money in general. I thus try to use question 6 to ask interviewees again about needs, motivations, and money, but this time indirectly and rather negatively (when would you stop?). I hope that this question will help interviewees to look at their filmmaking from a different angle.

### Data analysis

As mentioned above, my data analysis will follow the method of Analytic Induction. This decision will also influence the data collection process.

Analytic Induction was first introduced by Walter Znaniecki in 1934 and then later used and further theorised by, among others, Robinson (1951), Lindesmith (1952), Weinberg (1952) and Becker (1958). The research process that follows the method of Analytic Induction can be summed up by following steps (Robinson, 1951; Becker, 1958; Bogdan and Biklen, 1982; Ragin, 1994):

1. Formulate a rough definition of the phenomenon of interest.
2. Collect data.
3. Develop an initial theory and categories,<sup>10</sup> explaining the phenomenon.
4. Collect more data: Challenge the initial theory and categories by looking for deviant and negative cases.

---

<sup>7</sup> This follows, in part, the findings of Koestler (1964) and Becker (1982), who contend that in order to get made, art requires knowledge, ability, and a will to create (see above).

<sup>8</sup> Because of its centrality in existing discussions, I decided to give the monetary aspect its own Grand Tour question.

<sup>9</sup> Achievement, recognition, work itself, responsibility, and advancement.

5. Redefine initial theory and categories.
6. Repeat steps 4 and 5 until new data does not add new knowledge; i.e., new data perfectly fits theory and categories without any redefinition.

Analytic Induction follows what Glaser and Strauss (1967) termed the constant comparative model. This involves comparing data and theory/categories and using this assessment to constantly improve theory/categories. Analytic Induction can thus be used to theorise as well as to test the emerging theory by seeking contrary evidence. Looking for deviant and negative cases is hence especially important, as those cases either lead to a narrowing down of categories, a redefining of the theory, or a complete rejection of the theory. Robinson (1951) illustrated this necessity in his critical account of the method (see Table 1).

**Table 1**

	Phenomenon occurs	Phenomenon doesn't occur
Conditions present	X	?
Conditions not present	0	X

(from Robinson, 1951)

Researchers who first used Analytic Induction only looked at cases in which the phenomenon occurred. Using the example of my own research, this would mean that I only look at independent filmmakers. Researches then found conditions that were present in all positive cases. (In my example, to speculate, a possible condition may be that all independent filmmakers attended a formal camera workshop to learn the craft of cinematography.) Since all of the positive cases had the condition present, it was assumed that all cases in which the condition was not present had to be negative. (In my example, if someone did not attend a formal camera workshop, this person would not become an independent filmmaker).

Robinson (1951) argued that a researcher who only looks at positive cases can never be sure if the conditions were not also present in some negative cases, which would then completely refute the developed theory. (In my example, if some people who attended a formal camera course did not become independent filmmakers, the completion of a camera course would not predict a future as an independent filmmaker.) Robinson thus insists on examining negative cases using quantitative methods in order to verify findings. Other researchers replied to Robinson's critique by stating that researchers using Analytic Induction are required to look at deviant (Lindesmith, 1952; Weinberg, 1952; Becker, 1958) and even negative (Becker 1973) cases<sup>11</sup> in order to challenge their theory. Contrary to Robinson, Becker's negative cases are not randomly chosen ones in which the phenomenon does not occur, but cases that are chosen because the conditions are present but the phenomenon does not occur.

Interactionist thinking follows Becker's suggestion, for two reasons. First, quantitative data does not emerge in a social interaction but is heavily influenced by the researcher's own ideas. Other than qualitative methods, quantitative data collection is not open to the suggestions of the subjects studied, but relies solely on the researcher's preconceptions. Secondly, positivism puts understanding in the centre of an investigation rather than the quest for predictability. Human behaviour is far too complex for any kind of prediction (Bogdan and Biklen 1982). Qualitative research thus tries to describe and understand human behaviour rather than predict it.

What needs to be stressed at this point is a) that while Analytic Induction does try to test its theory by looking at deviant and negative cases, it nevertheless b) always remains a qualitative method that tries to explore new ground and bases its theory on the description, and possible explanation, of a phenomenon.

---

<sup>10</sup> Categories are needed if a theory does not fit all positive cases.

<sup>11</sup> The difference between deviant and negative cases is that deviant cases are cases in which the phenomenon occurs but other conditions apply or the (until then assumed) explanatory condition does not apply. Negative cases are cases in which the phenomenon does not occur.

Following the method of Analytic Induction and the constant comparative method, I will constantly compare theory and data. I will analyse initial interviews and develop an initial theory. I will then conduct additional interviews to test my initial theory. The findings of the additional interviews will then lead to a possible redefinition of the theory and/or the introduction of types (categories) of filmmakers, in case the emerging theory does not fit all filmmakers. I will stop interviewing people if new data does not add anything to my theory or does not demand the introduction of new categories.

### **Sample and representation**

My sample will include independent filmmakers from the UK. I picked my first interviewees from contacts I was given by my supervisor. From there, I will use a snowballing effect, asking the filmmakers I interview about further contacts. In accord with Analytic Induction, I will use "theoretical sampling," in contrast to "random sampling," meaning that I will look for deviant and negative cases that are likely to challenge emerging theories.

As there is no complete list or register of all filmmakers in the UK, I will not be able to have a random draw of interviewees. I will therefore employ the method of theoretical sampling. Accordingly, my sample will not be representative of the whole population of filmmakers in the UK, and my study will not fulfil the requirements for representation of a bigger population.

### **Discussing other methods**

It makes sense for me to use a qualitative method because it helps in the exploration of a new field and tries to develop a theory. Since no theory exists in my case, theory testing would mean developing a random theory.

I believe that open interviews and subsequent Analytical Induction are the best methods for developing a definition of sustainable independent filmmaking. All other methods have several limitations:

- I will not use participatory observations for two reasons:
  - They do not allow me to make assumptions about the past.
  - Following an individual from early childhood to life as a filmmaker is not feasible since it would take too long and, more importantly, I could not be sure that my subject would turn out to be a filmmaker.
- I will not use quantitative forms of research (survey, questionnaire) because they would allow too much room to my own preconceptions and might thus lead the research results to reflect my prejudices. This is because I would need to define indicators for categories of interest prior to data collection. This would limit my research and prohibit openness for new concepts or categories.
  - For the same reason, I decided against standardised interviews in favour of open interviews.
- Focus groups might have been an option. However, since my belief is that individual circumstances of becoming a filmmaker are very personal and differ from case to case, I saw the danger that a focus group might silence some personal histories and hence distort results in favour of the louder voices.

### **Issues with the chosen method**

Before describing the results of my research, I want to draw attention to some theoretical issues attendant to the method I have chosen.

The issue that concerns me most is that I can never be entirely sure if people are telling the truth. However, first, there is no reason for me to believe that interviewees would agree to be

interviewed for one or two hours and then deliberately lie to me. Second, I inform interviewees about my overall research aim, which is to find ways for independent filmmakers to use the internet in order to sustain their work. This makes me a person who tries to help independent filmmakers. I thus present myself as a “friend” into the interviewees, not as someone who is researching a debateable topic. Third, I designed my Grand Tour questions in such a way that the interview comes back to a topic and tries to cover it from a slightly different angle. This allows me to verify earlier answers.

Another major concern is the topic of money. Talking about money is always delicate. I fear that filmmakers might give socially desired answers—for example, that their filmmaking money is not motivated or enabled by money but only by their own desire to create. My strategy to avoid this dilemma is twofold:

1. As stated above, I will ask the questions about motivations and needs from slightly different angles. This way, I hope to uncover contradictions.
2. I will analyse the whole interview in order to develop a theory. This means that I will be able to compare the answers to questions related to motivation and needs with actual decisions from the filmmakers’ career paths. This, too, might uncover contradictions.

McCracken (1988) raises the issue that the researcher is always part of a culture. In my case, I am not an independent filmmaker, but clearly someone who went through the same formal education as most aspiring filmmakers. This gives me the opportunity for great insight into the right questions to ask, but can also lead to blindness concerning some problems. It is a challenge to try to use the advantages of insight and to try to work around the dangers of familiarity. Acknowledging and being aware of this dilemma are the best ways of dealing with it.

Becker (1982) made me think that my focus on independent filmmakers might be wrong and that I should instead understand the filmmaking process as collaboration. Following this thought, I should also interview all the collaborators (or at least the most important ones)—actors, producers, editors, camera and sound operators, etc. This is because if we want to sustain the filmmaking sector, we also need to sustain crewmembers. However, I believe that those supporting crewmembers often share the needs of the filmmaker. More importantly, I see the filmmaker as the driving force behind the making of a film. Also, it is possible to make films without any supporting crew. I hence limit my sample to filmmakers.

I am also aware that asking filmmakers about the past might distort answers. They might talk about the past in the way that they perceive it now. They might have forgotten things. They might turn randomness into a linear process. I will not be able to know how much the answers/stories are distorted and what might distort them. But I shall be aware of such possible distortion.

Finally, the circumstances in which art gets made, distributed, and appreciated must support sustainability (or rather the conditions needed to create it). In my research, I am arguing from the point of (more or less) free Western societies, knowing that even here there are huge differences in state or market support for filmmakers. I want to point out that I am not in a position to make assumptions about sustainable filmmaking in cultures other than my own.

## **Research results**

As mentioned above, Analytic Induction requires the researcher to constantly move between data collection and analysis. I decided to collect and analyse data in phases. For the purposes of this study, a phase will consist of a number of interviews (data collection) followed by an analysis of these interviews. This will then lead to an initial theory, possible categories of independent filmmakers, and a redefinition or verification of both. Results from a phase will influence the sampling and areas of interest of subsequent phases.

## Phase I

Phase I was composed of three interviews and their subsequent analysis. Interviews and their topics followed the description above. All interviews were tape recorded, transcribed, and analysed.

### *Career paths*

Becoming an independent filmmaker is usually dependent on two levels: arts and craftsmanship. On both levels, it needs discovery and exploration in a formal and informal way. Money does not seem to play a decisive role in the career path of an independent filmmaker.

The first step in becoming an independent filmmaker is discovery, which appears to take place on two levels. First, it requires exposure—not just to films in general but to films that one wants to follow or copy in style, storytelling, approach, and so forth. I will call this a filmic role model. Second, it necessitates a demystification of the filmic process. While a filmic role model is important as a guide for the creative part (film as art), demystification of the filmmaking is an important step for understanding the mechanical part (the craft) of filmmaking. This demystification usually takes place by playing around with the family's or a friend's camcorder:

Our family had a little 8mm camera, when I was a boy. Like many filmmakers, that's where the interest kicked in.

I was sort of interested in video cameras and people had just a camcorder. So I've always wanted to use it.

Regarding the filmic role model:

I was inspired in many ways by the innovative television coming out of the UK in the 60s and the 70s.

I saw a documentary called *The Revolution Must Not Be Televised* [sic]. . . . It's been criticised for being one-sided but it was kind of like "Wow, that's the kind of journalism I'd like to be doing."

It's the same reason why people start bands. It's because they like a certain type of music and they want to emulate that in some way. And I think that is basically how it started with me—probably quite naively, as a child, [I] just want[ed] to make a film that was like the films that I've been watching.

As a second step, it then seems to be necessary to receive formal education. A film course is important to further demystify the technical process of filmmaking and the film industry in general, and to enable the filmmaker to meet people and widen her knowledge of the film form.

So the course really opened my eyes to what was possible in independent and low-budget filmmaking.

Nonlinear edit packages, like Avid, came along, which I only came into contact with when I was at university back in the late 90s.

Once I started this course, I exposed myself to more filmmakers that I normally wouldn't even have heard of outside of the course.

After receiving formal education, the careers of independent filmmakers vary widely. However, what they have in common are persistence and commitment. They remain in or at the periphery of filmmaking and have a constant filmic output, which includes their own films (usually financed with their own money) as well as films for broadcasters. They work as crewmembers on commercial films or simply expose themselves to more films. This "doing" and "watching" has a number of effects. It further trains the filmmaker, allows her to network,

obtain insider tips, locate new work opportunities, and expose herself and her work to potential employers or funders.

I did another 5-minute thing about a politician trying to get elected, which wasn't as successful. But it was still good, just to be making a film.

But one of the main reasons why I did the film festival was to, again, going back to exposing myself to as much as possible.

I delivered by hand a letter to 127 different postproduction companies.

And of course the only way to get in is to be bloody thick-skinned and say, "Fuck you, I'm not going to give up." Which, luckily, I didn't.

**Table 2**

Career paths of independent filmmakers <sup>12</sup>						
IV	Filmic role model	Demystification	Formal education (arts)	Formal education (crafts)	Formal education (industry)	Constant filmic output
1	Yes	Yes	Yes	Yes	Yes	Yes
2	Yes	Yes	Yes	Yes	Yes	Yes
3	Yes	Yes	Yes	Yes	Yes	Yes

The results of phase I verify Becker's (1973) sequential model (see Table 2). They further add to my initial speculation<sup>13</sup> developing an interest in making films requires a filmic role model and demystification of the filmmaking process. Learning about the art and craft of filmmaking and the film industry takes place in a formal setting. Having a filmic output marks one's emergence as filmmaker.

### Needs

The results for the questions related to the enabling factors of independent filmmaking seemed, at first glance, very diverse and individual. At a second reading, however, it became apparent that the themes were remarkably analogous. They differed only in their relevance to the individual interviewees; for some, they were needs, for others wants.<sup>14</sup>

Central needs were time, people, equipment, and money. Motivations were exclusively intrinsic.

*Time* is necessary for the making of a film. However, the more production time is needed, the more potential there is for the film to come into conflict with income generation. All filmmakers interviewed needed time away from their day jobs (at the weekend, during the holidays, or after work) in order to make films. Having a job that is not too stressful and/or allows for managing one's time is essential for all of the three filmmakers interviewed.<sup>15</sup>

<sup>12</sup> The results summarised in Table 2 will possibly change during the coming phases of my investigation. They are meant to underline similarities. In subsequent phases, I will be able to test those findings. A "no" will then mean that a feature or stage is not needed in the career of an independent filmmaker or the making of independent films.

<sup>13</sup> To remind to reader: I speculated that a sequential career path of an independent filmmaker consists of developing an interest in filmmaking, wanting to make films, learning about filmmaking and, finally, making films.

<sup>14</sup> I divided the importance of themes into "needs" and "wants," with needs being absolutely necessary and wants being wishes that a filmmaker might hope for or work towards. For the purpose of this study, however, it makes sense to simplify the matter and join all of them under the heading "needs."

<sup>15</sup> The need for time also indirectly indicates a need for money. Everybody needs to survive. Not being able to live off of one's films thus entails reliance on other employment and making films during one's spare time. However, in my view money need not become the preeminent requirement for filmmakers as a consequence. All three interviewees were "part-time" filmmakers and all seemed to be happy with this status. This is not to say that they would not take jobs as full-time



Because of the sheer intensity of that work, it's very difficult to make your own film at the same time.

I work for myself. I can juggle my time and fit in filming around that.

*People* are a central need for all filmmakers. They are needed for many reasons—for example, as sources of inspiration and as colleagues. Amongst my interviewees, the latter reason was especially prevalent:

I truly, truly, truly would not have made it without the cocoon that was built around me. [T]he producer put me with a cameraman, who did a great job. A whole team around me who were there to support and create the film.

Another key point, again, meeting the other guys from the group—just basically becoming friends with them, realising that we all had similar interests, realising that we all can work together fairly easily.

But when going back to the course, one of the good things was that you met like-minded individuals. It's hard doing everything on your own.

It was interesting to see that all interviewed filmmakers had access to *equipment* when making their first film outside of formal education. They either bought their own equipment or their occupations put them into a position in which they had access to free equipment. All three thus said that they obviously need equipment to make a film, but also that having equipment at their disposal makes their filmmaking much easier, and thus helps enable it.

The interviews showed that the need for *money* is by no means as straightforward as suggested by Parker (2002). All three interviewees have different views regarding money and independent filmmaking. These views range from needing to raise production costs before making a film to money being a distant wish. A first theory is that established filmmakers, and possibly filmmakers with family commitments, do need to justify their time. They hence do not spend money on making a film, but ensure that they have raised production costs prior to making the film. Early career filmmakers, on the contrary, seem to take bigger risks and invest their own money in the making of a film. They hope that future employment or future funding will result from their investments.

There is another difference in the approach to money between the two early-career filmmakers I interviewed in phase I. When asked if they would continue making films if they would not earn money from them, one of them said that he would continue making films regardless of the monetary situation, while the other one said that he would need to justify his time and could not spend a lot of time on something that does not have a monetary return.

I'm hoping that [a] documentary might, in the long run, earn money or even if you get some exposure through a documentary. Money does matter. You can't really dedicate a lot of time and effort to something—especially if you got kids. I think it can be a secondary motivating factor—like it probably is for me.

You know, you just can't keep doing something for nothing. If you have some recognition, even if there wasn't a lot of money from [the film], it might lead to more commercial work, because people trust you or they like your style or whatever.

---

filmmakers. They would. More important to my research is their desire to work as part-time filmmakers as long as they have the time.

The first quote seems to imply that family commitments, or the prospect of them, turn money into a need for an independent filmmaker. The second quote, however, indicates that monetary needs may be satisfied both directly (through the film) and indirectly.

When asked about their motivations to make films, the reasons given were mostly *intrinsic* in nature:

[I]t was an artistic outlet.

I had this vision of just going out there fearlessly, telling the truth, telling untold stories.

The second quote indicates that audiences are an important factor. Indeed, the interviews reveal that both audience and feedback worked as motivators. Getting *recognition* was seen as something encouraging and reassuring, an indication that one is doing something of value.

And basically . . . what was nice was that . . . people appreciated it. It won a prize at Marseilles . . . It got shown around a lot. I got invited to various places.

I was quite surprised to see something that I've done entirely on my own having a good impact on people. That was really encouraging.

Well, it would be quite depressing if I would make film after film and none of them got accepted into any film festival and no one would want to watch them.

Payment can also work as reassurance and encouragement:

It was good to be spending time on something, knowing that you are getting paid. It wasn't just a labour of love.

Herzberg et al.'s (1959) factors of *achievement*<sup>16</sup> and *advancement*<sup>17</sup> could be confirmed as being motivators for independent filmmakers. So is the process of *filmmaking itself*. All of those factors motivated the interviewees. None of them identified responsibility as a motivating factor. This possibly has to do with the fact that filmmakers are ultimately responsible for their own work—with the exception of dependent workers (such as those studied by Herzberg et al).

A central theme of all of the interviews was the need for time, people, and equipment, along with the motivators of recognition/reassurance, achievement, advancement, and the work itself. Money and a filmmaker's view of it seem to be important differentiators. I will need to further investigate this in subsequent phases of the enquiry.

Phase I thus verifies Becker's (1982) and Koestler's (1964) findings; it also partially confirms Herzberg et al.'s (1959) results. Money—as emphasised by Parker (2002)—is, however, only one of a number of needs. It is neither the only nor the central need for the independent filmmakers interviewed. However, the need for money seems to be a central differentiator between individual filmmakers (see Table 3).

**Table 3**

Needs/motivations of independent filmmakers								
IV	Time	Meeting the "right"	Access to equipment	Recognition, reassurance	Achievement	Advancement	Work itself	Money

<sup>16</sup> Filmmakers were motivated by the possibility of achieving their goals with their films, which included career goals, the opportunity to "tell the truth," and simply desire to have an artistic outlet.

<sup>17</sup> The two early-career filmmakers were especially motivated by possible advancement into full-time, commercial filmmaking.

		people to work with						
1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	In future
3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No

### *Methodological considerations after phase I*

It was very hard for the interviewees to answer the questions about their motivations to be independent filmmakers. That said, all of them gave reasons. Their hesitations might thus indicate further intrinsic motivations, which are more emotional and theoretical than literal and empirical. The latter can easily be verbalised because of their occurrence in every day life, whereas it is harder to articulate emotions.

### *Phase II*

The goal of phase II will be to challenge my findings from phase I by actively looking for deviant cases. For example, two of my interviewees were filmmakers who did not receive formal education in film form and film craft. As the interviews of phase II have yet to be analysed, I am not able to add to the findings of phase I at this point.

### **Conclusion and road map**

I am currently in the process of conducting and analysing interviews in phase II. Phase I confirmed my speculation that there is more to sustaining independent filmmaking than pure economics. Although economics seems to play an important role, this aspect is neither straightforward nor singular. It rather looks as if money can be everything from an enabler to a secondary motivator to a distant want.

The results from phase I reaffirm that it is right to qualitatively interview independent filmmakers. As far as I can judge at this point, the interviews were able to widen my understanding of and challenge my pre-existing ideas about the careers and needs of independent filmmakers.

With regard to my overall PhD research project, the definition of the term *independent filmmaking* is, as I mentioned in the beginning of this paper, only a first, albeit necessary, step. Once such a definition is satisfactorily developed, I will be able to apply it theoretically to an internet environment. In other words, once I know what filmmakers need and want, I will be able to approach the question, "Is the internet able to provide independent filmmakers with these needs and wants?" I am planning to answer this question by a) looking at examples (case studies) of how independent filmmakers (and other artists) used the internet in the past and attempted to learn from those experiences, and b) by collecting additional data through experimenting with my own films.

## References

- Amabile, T. (1996) *Creativity in Context: Update to the Social Psychology of Creativity*. Oxford: Westview Press.
- Broadcasters' Audience Research Board (2009) 'Monthly multi-channel viewing summary' (Electronic database). Available at:  
[http://www.barb.co.uk/report/monthlyViewingSummary?\\_s=4](http://www.barb.co.uk/report/monthlyViewingSummary?_s=4).
- Becker, H.S. (1953) 'Becoming a marihuana user,' *The American Journal of Sociology* 59 (3), pp. 235–42.
- Becker, H.S. (1958) 'Problems of inference and proof in participant observation.' *American Sociological Review* 23, pp. 652–60.
- Becker, H.S. (1973) *Outsiders: Studies in the Sociology of Deviance, 2<sup>nd</sup> edition*. New York: Free Press.
- Becker, H.S. (1982) *Art Worlds*. Berkeley: University of California Press.
- Bogdan, R.C. and Biklen, S.K. (1982) *Qualitative Research for Education*. Boston: Allyn & Bacon.
- Bouchard, C. (Writer) (2009). *The Hunt for Gollum* (Film). United Kingdom.
- Centers, R. and Bugental, D. (1966) 'Intrinsic and extrinsic job motivations among different segments of the working population,' *Journal of Applied Psychology*, 50 (3), 193–97.
- Cramer, W. (2008). 'Sprechblasen im Ungleichgewicht.' Available at:  
<http://www.spiegel.de/wissenschaft/natur/0,1518,559142-2,00.html>
- Elsaesser, T. (2005) *European Cinema: Face to Face with Hollywood*. Amsterdam: Amsterdam University Press.
- Gardner, E. (2007) *Developing an Environmental Strategy for UK Film*. Available at:  
[http://www.ukfilmcouncil.org.uk/media/pdf/2/t/Developing\\_an\\_Environmental\\_Strategy\\_for\\_UKFC.pdf](http://www.ukfilmcouncil.org.uk/media/pdf/2/t/Developing_an_Environmental_Strategy_for_UKFC.pdf).
- Geylor, B. (Writer) (2009). *RIP: A Remix Manifesto* (Film). USA.
- Glaser, B. G. (1992) *Emergence Vs. Forcing: Basics of Grounded Theory Analysis*. Mill Valley, CA: Sociology Press.
- Glaser, B.G., and Strauss, A.L. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine Publishing.
- Herzberg, F., Mausner, B., and Snyderman, B.B. (1959) *The Motivation to Work, 2<sup>nd</sup> Edition*. New York: John Wiley & Sons.
- Independent Cinema Office (2001) 'Sustainable development for cinemas & film festivals.' Available at:  
[http://www.independentcinemaoffice.org.uk/training\\_2009\\_sustainabledevelopment.html](http://www.independentcinemaoffice.org.uk/training_2009_sustainabledevelopment.html).
- Institute for International Film Financing (2009) 'Overview.' Available at:  
<http://www.filmfinancing.org/aboutus.html>.
- Kaowthumrong, P. (2009) 'Colorado's film economy is on the rise, thanks to "sustainable" independent effort: House Bill 1010 seeks to establish new office of film, television and media,' *ColoradoBiz*. Available at: <http://www.cobizmag.com/articles/colorados-film-economy-is-on-the-rise-thanks-to-sustainable-independent-eff/>.
- Kluge, S. (Writer) (2008) *The Last Drug* (Film). Germany.
- Koestler, A. (1964) *The Act of Creation*. New York: Macmillan.
- Lindesmith, A.R. (1952) 'Two comments on W.S. Robinson's "The logical structure of analytic induction,"' *American Sociological Review* 17 (4), pp. 492–93.
- Macnab, G. (2009) 'Theatres of opportunity,' *Screen Daily*. Available at:  
<http://www.screendaily.com/territories/uk-ireland/theatres-of-opportunity/5002628.article>.
- McCracken, G.D. (1988) *The Long Interview*. Newbury Park, CA: Sage.

- O'Riordan, T.D. (1998) 'Indicators for sustainable development.' Paper presented at the European Commission (Environment and Climate Programme) Advanced Study Course, 5–12 July 1997, Delft, The Netherlands.
- Parker, A. (2002) 'Building a sustainable UK film industry.' A presentation to the UK film industry, 5 November 2002. Available at:  
<http://www.ukfilmcouncil.org.uk/media/pdf/q/r/BaSFI.pdf>.
- Popper, K. (1980) *The Logic of Scientific Discovery*. London: Unwin Hyman.
- ProjectX (2009) 'HD24P.' Available at:  
[http://www.elementalfilms.co.uk/projectx/download/px\\_hd24p.pdf](http://www.elementalfilms.co.uk/projectx/download/px_hd24p.pdf).
- Ragin, C.C. (1994) *Constructing Social Research: The Unity and Diversity of Method*. Thousand Oaks, CA: Pine Forge Press.
- Robinson, W.S. (1951) 'The logical structure of analytic induction,' *American Sociological Review* 16 (6), pp. 812–18.
- Weinberg, K.S. (1952) 'Two comments on W.S. Robinson's "The logical structure of analytic induction,"' *American Sociological Review* 17 (4), 493–94.

# Social Enterprise Applications in an Urban Facilities Management Setting: Research Methodological Perspectives

**Siti Uzairiah Mohd Tobi**, *University of Technology Malaysia; School of Built and Human Environment, University of Salford*, **Dilanthi Amaratunga**, *School of Built and Human Environment, University of Salford*

## Abstract

The credibility of research findings is an important aspect of any research success. This is influenced by the selection of an appropriate research methodology. Research methodology is the overall approach to be used in the research process from the theoretical underpinning to the collection and analysis of the data. In this context, the proposed paper will discuss the research methodology adopted for the study in developing an approach towards a new service delivery model that meets the needs of social enterprise philosophies in urban facilities management (urban FM) setting. It will further appraise the suitability of the research methodology for the study within the identified context. In doing so, the paper elaborates the theoretical underpinning related to research philosophies, strategies and techniques for data collection and analysis with details of appropriate justifications.

## Keywords:

New service delivery model, Research methodology, Social enterprise, Urban FM

## 1 Introduction

The increasing costs have been reflected in local authorities which carry the responsibility of facilities operation (Alexander and Brown, 2006; Ngowi and Mselle, 1998). The need to move away from the traditional service provider's approach is seen as a way of reducing running costs in facilities operations. Therefore within this context, this research considered taking an urban facilities management (urban FM) as a mechanism to develop a sustainable design to manage public facilities operations, by having an approach to develop a new service delivery model that meets the need of social enterprise. Since research is an activity that needs to be approached with both discipline and rigour (O'Leary, 2004) adoption of a suitable methodology is an important step for its credibility. Adding to that, Sutrisna (2007) shows that methodology takes the most time in research as the researcher attempts to place the work among the existing works on the topic, drawing on insights from a wide-range of literature reviews and developing an innovative angle on the topic. Accordingly, this paper identifies the research methodology adopted for the study, and examines its suitability by bringing forward the theoretical underpinning related to philosophies, strategies and data collection as well as methods analysis. At this point, this study has been based on theoretical and practical ideas obtained through comprehensive works of literature review.

In brief, this paper is organised into three sections: the first section describes the development of research objectives from the identified gaps in the literature through a brief review of related concepts. The second section portrays the existing research philosophies and methodologies, while highlighting the main facets of the arguments on their relative characteristics; whereas, in the final section, issues that need to be considered when selecting the most appropriate approach and research methods are outlined, while justifying the selected research strategy for a social enterprise applications in an urban FM setting.

## 2 Background to the study

### 2.1 *Need for a new service delivery model in public facilities operations*

The literature review has revealed that one of the major problems faced by local authorities in running the public facilities operations is the ever-increasing annual operational costs

(Alexander and Brown, 2006; McShane, 2006; Ngowi and Mselle, 1998). In seeking solutions to this problem, approaches such as privatisation and outsourcing have been applied. These approaches seem to be more efficient than most of the public services providers. However, there are arguments about the ways in which private companies should protect the public interest, while at the same time being profit-seeking organisations, as this could create a conflict of interest. Therefore, this study identified the need to move away from the traditional service provider's approach, as stated above. This could bring about a new concept for developing a sustainable design to maintain public facilities for community benefit. By considering sustainable factors in seeking new ways of delivering public services, the study, in the context of urban FM, could be regarded as having a new service delivery model to manage public facilities operations for urban sustainability.

The primary objectives are to explore the urban FM concept, as it is a part of the new FM alignment, and to understand the underlying philosophy of urban FM in relation to social enterprises principles, in order to understand the applicability of social enterprises principles as a new service delivery model in the urban FM setting. The target could be reached by exploring the variety of service delivery models of social enterprises that suit an urban FM setting. The conceptual arguments from the literature review centered around two main concepts; urban FM as a new FM direction, and its relation to Social Enterprise, with an emphasis on managing public facilities to be considered as an issue for this study. The result will lead into identifying the characteristics of such a model in order to develop the model to be applied in the Malaysian context.

As a developing country, Malaysia also faces the same problems in managing the public facilities operation in local authorities as those addressed above (Tobi, 2003; Tobi, 2006). The Malaysian Prime Minister urged the public and private sectors to come up with a more effective and efficient procedural framework to continuously improve the management of national assets and facilities. He further agreed to have an annual convention, the National Asset and Facility Management Convention (NAFAM), to oversee this issue, and the first inaugural NAFAM was held in August 2007. This provides a clear justification for the study in seeking its application to the Malaysian setting. Although the researcher decided to take urban FM (New FM alignment) as an approach to deal with this problem, nevertheless cultural issues and contextual factors need to be taken into consideration by practitioners. This is especially true for those who work or implement the practice outside the UK and USA; as Wong (2000) indicates, the cultural influence on FM practice in Asia in terms of standard, perception, and quality is significant. He points out that differences in the condition and specifications of facilities, standards, costs, and culture can affect FM practice, and also emphasises that FM practitioners need to be sensitive to cultural issues. These differences are equally applicable in the Malaysian setting, which seems to have lower standards compared to those of the UK and USA, due to economic capacity and cultural tolerance as well as local preference (Wong, 2000).

## 2.2 *Aim and Objectives*

The research aims to develop new service delivery model that meets the needs of social enterprise principles in an urban FM setting for managing community facilities operations within the Malaysian setting. To achieve this aim, the following objectives have been formulated:

- To explore urban FM principles within the concept of "new FM alignment"
- To understand the underlying philosophy of urban FM and its relationship with social enterprise principles
- To understand the applicability and characteristics of social enterprise principles as a new service delivery model in the urban FM setting
- To develop and test the applicability of such a model in the Malaysian setting in using it as an enabler in managing community facilities operations.

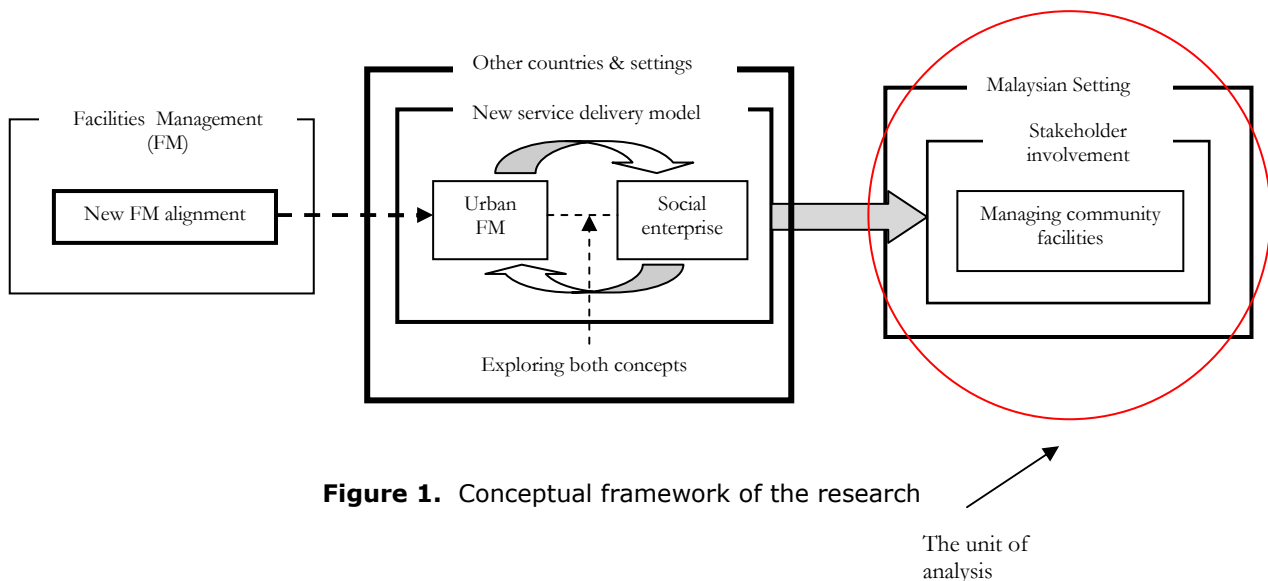
With an emphasis on new ways of delivering public facilities to be considered as an issue for this study, it is assumed that the relationship between urban FM and social enterprise would be able to provide a strong base to formulate an approach to develop the new organisational form that is sought. Currently, the form of social enterprise has not been established yet in the

Malaysian economic structure as compared to other regions in the world such as Britain, America, Europe, Australia, India and Hong Kong (EPU, 2001, EPU, 2008, Kasim AND Jayasooria, 2001). The research problem is formulated based on the premise that there is a lack of application of social enterprise practices in the Malaysian economic structure.

### 2.3 Conceptual Framework

Conceptual frameworks act as maps that give coherence to empirical inquiry and take different forms depending upon the research question or problem (Kaplan, 1964). A conceptual framework is able to explain, either graphically or in narrative form, the main things to be studied – the key focus, constructs or variables – and the presumed relationships among them (Miles and Huberman, 1994). It is a major step involved in the research process before proceeding to the next stages of the study, which are research approach and research techniques for collecting data as well as analysing it. The study aims to develop a new service delivery model that meets the needs of social enterprise principles in an urban FM setting for managing community facilities operations within the Malaysian setting.

The primary objective is to understand the applicability and characteristics of social enterprise principles as a new service delivery model, in the urban FM setting, by learning lessons from other countries in a different setting. Later, this objective will be further developed to test the applicability of such a model to the Malaysian setting in using it as an enabler in managing community facilities. In this context, the conceptual framework depicted in Figure 1, shows how the research problem is embedded within the scope of the study.



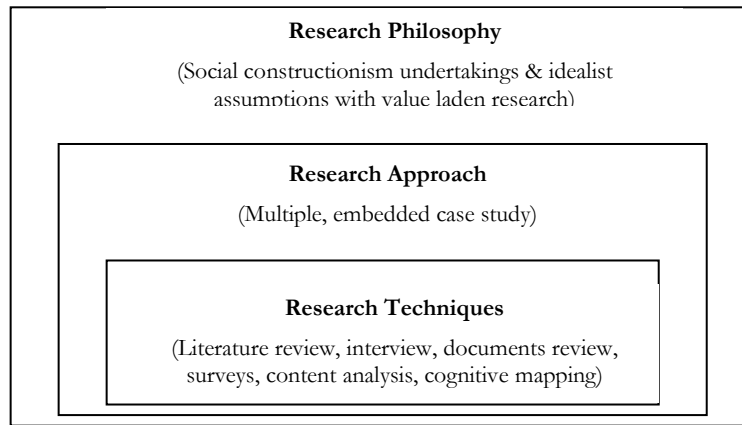
**Figure 1.** Conceptual framework of the research

### 3 Research Methodology

Research methodology refers to the procedural framework within which the research is undertaken, where the approach can be put into practice to help describe the research process (Remenyi et al., 2003; Collis & Hussey, 2003; Leedy, 1989). The researcher needs to understand each of the components in the research methodology to better understand the importance and relation of each component in order to coordinate the right flow of the research process. Towards achieving the aim of this research, it will follow the methodology of nested approach introduced by Kagioglou et al., (1998) as it helps to understand the assumptions and path dependencies of this research.

The following figure illustrates a summary on deployment of each element of nested approach methodology in this study.





**Figure 2.** A summary of applied research approach

**Source:** Nested Approach (Kagioglou et al., 1998, Kagioglou et al., 2000)

As this study deployed a nested approach research strategy by Kagioglou et. al (1998 & 2000), each layer of the nested approach will be discussed in the following section.

### 3.1 *Research Philosophy*

At the first layer of the nested approach, it explains the philosophy that is chosen with the appropriate assumption in order to classify: What is count to be reality in many ways to structure it as type of knowledge? (Ontology), how the knowledge of that reality may be established (Epistemology) and what values go into that knowledge (Axiology).

As a whole, research philosophy contains important assumptions that will underpin the research strategy and research methods chosen as part of that strategy (Saunders et al., 2007). This process helps the researcher to gain knowledge based on the empirical study, therefore it makes a contribution to the body of knowledge in an appropriate manner.

#### 3.1.1 *Epistemology undertaking*

Epistemology is about how the world has been viewed in reality. It is a general set of assumptions about how we acquire and accept knowledge about the world (Sexton, 2008). Based on the epistemological stance, Easterby-Smith et al. (2002) indicate the two contrasting ends of the philosophical traditions continuum about how social research should be conducted; positivism and constructionism (interpretivism).

In different ways, many authors and researchers (Belbase, 2007; Sexton, 2004, Creswell, 2003; Easterby-Smith et al., 2002; Remenyi et al., 2003; Amaratunga et al., 2002) agree that positivism is a philosophy which states that the only authentic knowledge is knowledge that is based on actual sense of experience. Such knowledge can only come from affirmation of theories through strict scientific method. In contrast, the constructivist view argues that knowledge and reality do not have an objective or absolute value or, at the least have no way of knowing this reality. Constructionism is more likely a view in philosophy according to which all knowledge is constructed in as much as it is contingent on connection, human perception and social experience.

It is important that the epistemological stance of the research needs to be looked at together with the ontological assumption and axiological purpose, as it is a part of the philosophical branch of knowledge. It will help the researcher to better understand the research that he/she undertakes even from an early stage in order to help clarify objectives, as this will make a significant contribution to the body of knowledge.

#### 3.1.2 *Ontology assumptions*

Ontology is an assumption that the researcher makes about the nature of reality (Sexton, 2008). It is a study of conceptions of reality and the nature of being. It seeks to describe or

posit the basic categories and relationships of being or existence to define entities and types of entities within its framework. As a philosophical subject, ontology deals with the precise utilization of words as descriptors of entities or realities. Sexton (2004) in his model of research approaches continuum shows that ontology could fall under realism or idealism of research knowledge. Ontological philosophy helps the researcher to place the research base on the knowledge of realism or idealism.

### 3.1.3 Axiology purposes

The axiology purpose is an assumption about the nature of values and the foundation of value judgements (Sexton, 2008). Axiology is also known as philosophical fields that depend crucially on notions of value, and sometimes it is held to lay the groundwork for these fields. The nature of value could be determined: either it is value free and unbiased, or it is value laden and biased. Further, these three assumptions will help to position the research within the philosophical continuum.

The position of the research paradigm for this study is summarised as follows:

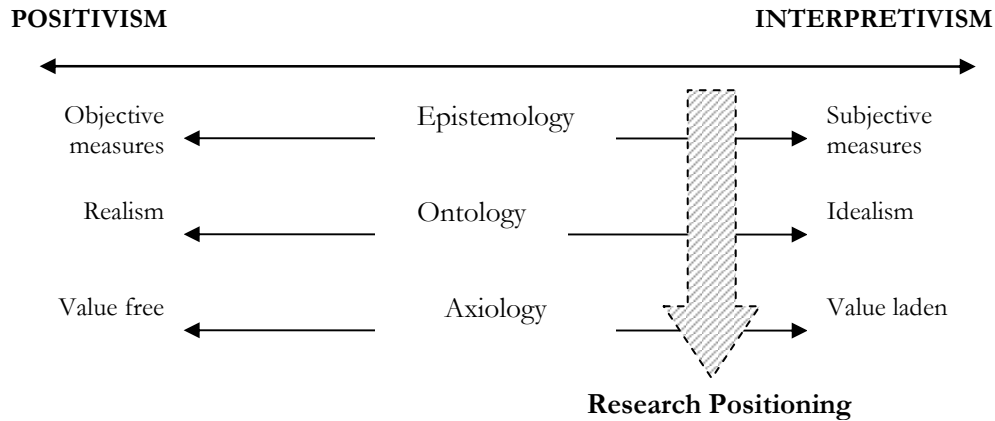
Research positioning towards interpretivism has been identified as the most appropriate research philosophy for this study. As the main research aim is to have an approach to develop such a model of new organisational form that complies with social enterprises development through urban FM setting for Malaysian application. Justification of the study is built around the two main concepts of urban FM and social enterprise.

With the emphasis on delivery of public facilities in this study, the interrelationship between those two main concepts will be able to provide a strong base for an approach to develop the new organisational form that is sought.

- *Epistemological* undertaking, as this study leans more towards interpretivism, the nature of this study is rooted in the notion of lived-world experience. This study is socially constructed based on knowledge gathered by examining the variety of service delivery models in social enterprise and urban FM.
- *Ontological* assumption, the study is largely a theory building attempt rather than a theory testing attempt, therefore it holds the ontological assumption that reality is not pre-determined, but socially constructed. Additionally, the research environment was not expected to be controlled and simplified with assumptions and hypothesis as in the deductive research approach used in positivist studies. Conversely, an inductive research approach is used with the intention of generating rich data to build up theories. The aim of this study is to construct and develop an approach for a new service delivery model that meets the need of social enterprise in an urban FM setting for public facilities operation.
- *Axiological* purpose, this study leans more towards the value laden as the research choices are determined by human beliefs and interests (Collis AND Hussey, 2003, Easterby-Smith et al., 2003). The phenomenon under study is interpreted within a context through direct interactions with local government, in order to seek the contextual factors to explore the applicability of such a model in Malaysian setting. Thus, the research environment cannot be controlled as the idea constructed is determined by human beliefs and interests. Accordingly, the study holds a value laden research choice.

Accordingly, Figure 3 below illustrates the positioning of research within the philosophical continuum within which it may fall pertaining to this study.

## Research Philosophies



**Figure 3.** Positioning the research within the philosophical continuum

In summary, the nature of this study is contributing to the establishment of a research philosophical base within the epistemological territory of social constructionism (interpretivism), ontological stand of idealism and axiological view of being value laden and therefore possibly biased, as it is determined by the experience and interpretations of the researcher. This philosophical standing of the research influences the selection of an appropriate research approach as described within the next section.

### 3.2 Research Approach

The research approach is the strategy taken towards data collection and analysis. It is also about the contributing knowledge claims and method (Creswell, 2003). It depends on the research inquiry and it could be a quantitative or qualitative inquiry in order to undertake the study. Those research approaches so much rely on the philosophical stances of the research being undertaken. Strategies of inquiry are associated with the research approach. Whether it is a quantitative, qualitative or mixed methods approach, it is a strategy that the researcher needs to be firm about. The following table (Table 7.2) shows alternative strategies of inquiry (Creswell, 2003).

**Table 1:** Alternative strategies of enquiry – Creswell (2003)

Quantitative	Qualitative	Mixed Methods
Experimental designs	Narratives	Sequential
Non-experimental designs, such as survey	Phenomenologies	Concurrent
	Ethnographies	Transformative
	Grounded Theory	
	Case studies	

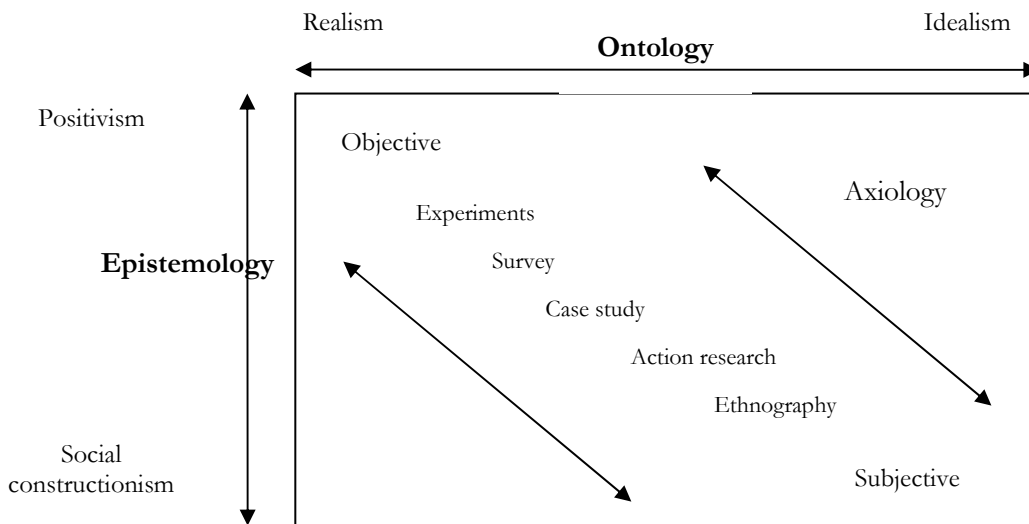
In addition, the method of data collection and handling could be differentiated between quantitative and qualitative inquiry. This picture is set out in Table 2 (Silverman, 1998) below.

**Table 2:** Two schools of science – Silverman (1998)

Approach	Concepts	Methods
<b>Positivism</b>	Social structure Social facts	Quantitative Hypothesis testing
<b>Interpretive science (phenomenological)</b>	Social construction Meanings	Qualitative Hypothesis generation

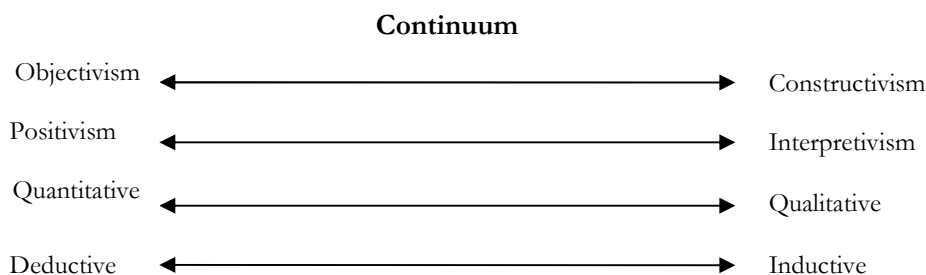
Strategies of inquiry are associated with the research approach. Whether it is a quantitative, qualitative or mixed methods approach it is a strategy that the researcher needs to be firm about. The researcher brings to the choice of a research design assumptions about knowledge claims (Creswell, 2003). These strategies of inquiry - quantitative, qualitative or mixed methods - contribute to the overall research approach.

Apart from that, Sexton (2004) identifies research approaches as a continuum, based on the ontological, epistemological and axiological stands of the study.



**Figure 4.** Research approaches continuum (Source: Sexton, 2004)

In different ways, Gill and Johnson (2002) as well as Sutrisna (2007) assert that the aforementioned research approaches can be differentiated and also be placed along the philosophical continuum depending on their emphasis on deductive or inductive research, degree of structure and the type of data they generate. Figure 5 below illustrates the philosophical continuum depending on their deductive or inductive research approach.



**Figure 5.** Research approaches continuum (Source: Sutrisna, 2007)

Moreover, Hart (2003) also outlines the difference between deductive and inductive research procedure. Table 4 shows the differentiation base of research procedures.

**Table 4:** The comparison of deductive and inductive procedure for research – Hart (2003)

Deductive procedure for research	Inductive procedure for research
<p>The researcher tests a theory</p> <p style="text-align: center;">↓</p> <p>Hypothesis or research questions are derived from the theory</p> <p style="text-align: center;">↓</p> <p>Concepts and variables are operationalised</p> <p style="text-align: center;">↓</p> <p>An instrument is used to measure the variables in the theory</p> <p style="text-align: center;">↓</p> <p>Verification of the hypothesis</p>	<p>The researcher gathers information and data</p> <p style="text-align: center;">↓</p> <p>Questions are asked about the phenomenon</p> <p style="text-align: center;">↓</p> <p>The data is classified and placed into categories</p> <p style="text-align: center;">↓</p> <p>Patterns are looked for in the data and potential theories are proposed</p> <p style="text-align: center;">↓</p> <p>Theories are tested and developed and patterns compared with other patterns and theories</p>

The deductive approach is more suited to theory testing than theory building and therefore will be better for fields where there is a strong body of accepted theory, compared to the inductive research approach, which is more suited to theory building than theory testing for fields where there is not a strong body of accepted theory. This is particularly useful for researchers undertaking a piece of reflective study on an issue in which they have the scope to make changes.

In conclusion, this study falls under the philosophical research continuum within the epistemological territory of social constructionism (interpretivism), the ontological stand of idealism and the axiological view of being value-laden and biased. Therefore, this disqualifies it from taking an experiment or survey as a research approach, since they are governed by positivism and realism in terms of research philosophy. Furthermore, the aim of the study is more descriptive and exploratory in nature, with 'what' and 'how' questions as guides for the study. Thus, an approach suited to exploration and description is required. Nevertheless, this study could have an element of an explanatory approach to better explain the phenomenon in the case study. In summary, taking all these into account, a qualitative approach with a case study design has been identified as an appropriate approach for this study.

#### **4 Research Design – Case study**

The aim of the study is more descriptive and exploratory in nature, with 'what' and 'how' questions as guides for the study. Thus, an approach suited to exploration and description is required. Nevertheless, this study could have an element of an explanatory approach to better explain the phenomenon in the case study. Added to that, Creswell (1998) suggests that the research questions 'how' or 'what' can be classified as belonging to a qualitative study, and indicate five traditions of research design within the qualitative inquiry, namely: case study, biography, phenomenology, grounded theory and ethnography. Therefore, taking all these into account, a qualitative approach with a case study design has been identified as an appropriate approach for this study.

The case study is an empirical enquiry that rises out of the need to understand complex phenomena, within a real life context, by undertaking in-depth data collection involving multiple sources of information (Creswell, 1998; Remenyi et al., 2003; Yin, 2003). In determining the choices between single versus multiple case studies, and holistic versus embedded unit of analysis (Yin, 2003), being either a critical, unique, representative, revelatory or longitudinal case provides the rationale to select the single case method over multiple cases. However, selecting multiple cases can further add to the distinct advantages of multiple sources of evidence and replication of findings.

This study is taking the qualitative inquiry of a 'what' question as the first stage to be explored, and a 'how' question as the second stage to be explained. This study demands an in-depth investigation as an approach to develop such a model, as well as having a breadth of perspectives in exploring the applicability of the model in a Malaysian context. Therefore, having a unit of analysis in each individual case will guide towards achieving the main research aim. This research does reflect the multiple-case designs as compared to a single case design. Therefore, the unit of analysis of this study, community facilities operations, is embedded in finding the new organisational form for a Malaysian setting. Accordingly this research will deploy a qualitative approach for data collection technique by interviewing two target groups in each case study; using multiple embedded case studies is a means for triangulation of data sources to give strong findings. These groups act as several units of analysis involved in understanding Malaysian contextual factors.

#### 4.1 Unit of analysis for case study

The unit of analysis is one of the major entities that will be analysed in the study undertaken. It is because the analysis that the researcher planned to do will determine what the unit involves, in order to investigate the issue being researched. In many areas of social science research, as in this study being undertaken, the hierarchies of analysis units have become particularly important and have spawned a whole area of statistical analysis sometimes referred to as 'hierarchical modelling' (Trochim, 2006). According to many researchers (Kervin, 1992, Remenyi et al., 2003, Trochim, 2006, Yin, 2003b) the unit of analysis of a case study can range from an individual, a group of people, to a process or relationships.

Henceforth this study principally concentrates upon the public facilities operation as the main issue. Therefore, in exploring the applicability of such a model it will look into different local authorities as a multiple-case study; nevertheless it will treat each of them as an individual case.

The following figure as shown in section 2.3 of the conceptual framework, illustrates how the unit of analysis of this research is built within the scope of the study.



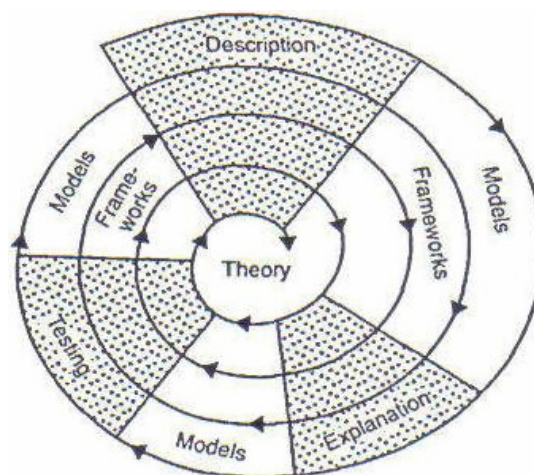
**Figure 6.** Unit of analysis

Accordingly, the case study approach will help in the exploratory phase and in the explanation of the causal process to clarify ways of thinking in relation to certain given problems or phenomena (Johannessen, 1997).

Therefore, this study involves an in-depth investigation into real life by looking at different theories, current practices, having expert opinions on theory building as well as exploring the applicability of the model in a Malaysian context. The next section will further discuss the case study process for theory building.

#### 4.2 Case study process for theory building

As noted by Meredith (1993), theory building attempts are possible to achieve through conceptual methods. The normal cycle of research from description to explanation to testing with continuing iteration could be described through the cycle as illustrated in Figure 7.4. Throughout this iterative process, descriptive models are expanded into explanatory frameworks which are tested against reality until they are eventually developed into theories as the research study builds upon them (Meredith, 1993).



**Figure 7.** The normal cycle of research for theory building

Adding to that, Yin (2003a) has mentioned that the purpose of case study for theory building structure comprises of explanatory and exploratory report writing in nature. Further, Lynham (2000) asserts that continuous refinement between theory and practice is also needed for effective theory building.

As discussed in the previous section, the multiple embedded case study approach is selected for this study. Accordingly, the researcher explored and recognised the social enterprise principle within the urban FM setting without controlling the diversity of service delivery models which are offered in social enterprises. The researcher took into account the various types of service delivery model that could be applicable, in order to be able to identify the characteristics of a new model for study in context. This is done by studying the inter-relationships between urban FM and social enterprise concepts.

The following section explains the research techniques that will be undertaken to conduct the case study.

#### 4.3 Research Techniques – Data collection and analysis

This section discusses the research techniques used in this research, mainly the methods of data collection and data analysis techniques to be used.

#### 4.3.1 Data collection procedure

This section includes a brief explanation on data collection methods, details of the selected case studies, details of the interviewees, the data collection plan and the expected preparation prior to field visits.

#### 4.3.2 Data collection methods

Methods of data collection are considered under the research techniques. The most important strategy to employ in data collection is to have multiple sources of evidence that can be converged with the same set of issues, which will later allow triangulation in order to construct validity for the study. Taking into consideration possible data to be collected in order to achieve the research objectives for this study, the data collection methods have been identified as follows:

- **Semi-structured interviews**

Semi-structured interviews are the main mode of data collection in finding the answers for most of the research questions. The interviews will be focusing on two target groups, at top level management in local councils, to seek their opinions as they represent the stakeholders in the issue under investigation (community facilities operations). The first target group will be heads of departments/senior officers, and the second target group will be the council members of several local councils in Malaysia. Along with the interview questions, this study will have a questionnaire survey which will be used as a complementary technique with the interview findings.

At each unit of analysis, the set of questions from the semi-structured interviews with survey questionnaires will be used to gather the information that is needed.

- **Document survey**

Document review will also be used as a supplementary technique as a means of triangulating the data collection techniques. Relevant local council documentation will be reviewed in managing the community facilities operations, as well as documentation that stated the involvement from others such as privatisation, outsourcing or joint venture projects related to community facilities operations. This method will be useful in understanding the context and the phenomenon of the case being studied.

#### 4.3.3 Data analysis methods

Data analysis is a component of research techniques and relates to the innermost layer of the nested diagram (Kagioglou et al., 1998, Kagioglou et al., 2000). As Burns (2000) indicates, the purpose of analysing the data is to find meaning in the data, and this is done by systematically arranging and presenting the information. However, in making the data collected become meaningful to the study, Easterby-Smith et al. (2002) deduced that, the analysis requires both a clear explanation of how the analysis was done, and a demonstration of how the raw data was transformed into a meaningful conclusion. There are several techniques that can be utilised to analyse data in case studies to improve the rigour in analysis, principally depending on whether qualitative or quantitative data has been collected. The principal data collected for this study, during pilot interviews and the explanatory phase of the case study, are qualitative in nature. Meanwhile, quantitative data, collected from the questionnaires survey as part of the exploratory phase, will be used to complement the interview findings.

The semi-structured interviews will be analysed using content analysis with a word-based and code-based approach. According to Krippendorff (2004) content analysis can range from the simplest form of word count to thematic analysis or conceptual analysis.



Further, this study deploys code-based qualitative content analysis and cognitive mapping as the techniques for analysing qualitative data, as the cognitive mapping technique counterbalances the weaknesses of the data display in code-based content analysis. Thereby, content analysis is used to identify concepts by developing codes, whereas cognitive mapping is used to explore relationships amongst concepts by illustrating the visual presentation. Consequently, cognitive mapping will be used as a technique to help bridge the gap between raw data and theory building. Finally, the questionnaire survey results will be entered onto spreadsheets using Microsoft Excel software. Subsequently, the data sheets will be transferred to SPSS software for descriptive statistical analysis.

#### 4.4 Methodological measures

This section identifies the design tests that are adopted in order to judge the quality of the research design in terms of its validity and reliability. Four widely used tests and the recommended case study tactics, together with the respective phase of research when each tactic is to be used, are outlined in Table 4 on the next page.

**Table 4:** Case study tactics for four design tests (adopted from Yin 2003)

Tests	Case Study Tactic	Phase of research in which tactic occurs
Construct validity	Multiple sources of evidence & establish chain of evidence to converge	Data collection
Internal validity	Pattern matching	Data analysis
External validity	Replication logic in multiple-case studies	Research design
Reliability	Case study protocol and database	Data collection

Construct validity and reliability are tested during the data collection stage. In order to satisfy the construct validity, the researcher needs to develop a sufficiently operational set of measures by using multiple sources of evidence in collecting data (Yin, 2003). The case study findings or conclusions will be much more convincing and accurate if based on several different sources of information. This will be achieved by using theory triangulation and methodological triangulation. The theory of triangulation will be used to get different perspectives on the same data set. Therefore, the identification of stakeholders involved in community facilities operations towards the new model that is sought would come from two target groups in local councils. Although in the same local council, these two target groups are distinct; the first group represents a public sector that is currently running the community facilities operations and holds assets ownership; the second group represents the community as the council members are chosen by the community to take care of their interests and might know what the community needs. This research will only try to focus on decision maker level/top management in order to gain an insight or understanding of the new model that will be developed. Furthermore, the semi-structured interviews as a qualitative method for data collection will be used together with survey questionnaires in order to seek opinions in developing the new model, as a complement to the interview findings which ensures the methodological triangulation. Moreover, the principle of case study protocol and developing a case study database will be employed to address the reliability of the design test and to demonstrate that the operations of the study can be repeated with the same results.

Internal validity refers to establishing the causal relationships (explanatory), whereby certain conditions are shown to lead to other conditions, as distinguished from spurious relationships (Yin, 2003). This is tested during data analysis stage by using pattern matching logic; a technique which compares the theories and observed data (Trochim, 1989; Yin, 2003). Accordingly, this research will compare the data collected through expert interviews and semi-structured interviews with the theoretically predicted one. If the patterns coincide, the results can help a case study to strengthen its internal validity.

External validity means establishing the domain to which a study's findings can be generalised by using replication logic, and is to be tested during the research design stage. The findings from one case study will be replicated by conducting a second, third and even more case studies. Therefore, using replication logic in multiple-case studies satisfies this external validity by using literal replication (within case analysis) and theoretical replication (cross-case analysis) to have the result more robustly.

## 5 Conclusion

Clear definition and design of a research strategy is a fundamental and necessary requirement for an empirical study in any research field. This paper has discussed thoroughly an appropriate research strategy to come out with an approach to develop a new service delivery model that meets the need of social enterprise principles in an urban FM setting. However, it could not be taken that the aforementioned research methodology is the only suitable design for similar research undertaken, which has been discussed within the paper; the suitability will depend on the aim, objectives and questions of the research being studied.

## References

- Alexander, K. and Brown, M. (2006) Community -based facilities management. *Journal of Facilities Management*, vol.24, (nos.7/8): pp.250-268.
- Amaratunga, D., et al. (2002) Quantitative and qualitative research in the built environment: application of "mixed" research approach. *Work Study*, 51, (1): 17-31.
- Belbase, S. (2007) *Learning and Teaching Mathematics from Traditional to Constructivism: a Portrayal of Pedagogical Metamorphosis*. Thesis Kathmandu University.
- Burns, R. (2000) *Research Introduction to Research Methods*. London: Sage Publications Ltd.
- Collis, J. and Hussey, R. (2003) *Business Research: A practical guide for undergraduate and postgraduate students*. 2nd ed. New York: Palgrave Macmillan.
- Creswell, J. W. (1998) *Qualitative Inquiry and Research Design-Choosing Among Five Traditions*. London: SAGE Publications.
- Creswell, J. W. (2003) *Research Design-Qualitative, Quantitative, and Mixed Methods Approaches*. 2nd ed. London: SAGE Publications.
- Easterby-Smith, M., et al. (2003) *Management Research: An Introduction*. 2nd ed. London: SAGE publications.
- Epu (2001) Eighth Malaysian Plan 2001-2005. *In: Malaysia: Economic Planning Unit, Prime Minister's Department, Malaysia*.
- Epu (2008) *Ekonomi Malaysia: Fakta dan Maklumat*. *In: Malaysia: Unit Perancang Ekonomi, Jabatan Perdana Menteri Malaysia*.
- Gill, J. and Johnson, P. (2002) *Research methods for managers*. 3rd ed. London: SAGE publications.
- Hart, C. (2003) *Doing a Literature Review-Releasing the Social Science Research Imagination*. London: SAGE Publications.
- Johannessen, J.-A. (1997) Aspects of causal processes in social systems: Discussions of methodology. *Kybernetes*, 26, (1): pp.30-51.
- Kagioglou, M., et al. (1998) Generic Design and Construction Process Protocol: final report. *In: The University of Salford*.
- Kagioglou, M., et al. (2000) Rethinking construction: the Generic Design and Construction Process Protocol. *Engineering, Construction and Architectural Management*, 7, (1): pp.141-153.
- Kaplan, A. (1964) *The conduct of Inquiry: Methodology for Behavioral Science*. Scranton, PA: Chandler Publishing Co.
- Kasim, M. Y. and Jayasooria, D. (2001) Informal Economy, Micro Finance and Non-Governmental Organisations in Malaysia. *Humanomics*, 17, (No.1/2): pp:134-140.

- Kervin, J. B. (1992) *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Krippendorff, K. (2004) *Content Analysis: an introduction to its methodology*. 2nd ed. California and London: SAGE.
- Leedy, P. D. (1989) *Practical Research-Planning and Design*. New York: Macmillan Publishing Company.
- Lynham, S. A. (2000) Theory building in the human resource development profession. *Human Resource Development Quarterly*, 11, (2): pp. 159-178.
- Mcshane, I. (2006) Community facilities, community building and local government-an Australian perspective. *Facilities*, vol.24, (nos.7/8): pp.269-279.
- Meredith, J. (1993) Theory Building through Conceptual Methods. *International Journal of Operations and Production Management*, 13, (5): pp.3-11.
- Miles, M. B. and Huberman, A. M. (1994) *Qualitative data analysis: An expanded sourcebook* 2nd ed. Thousand Oaks, CA: Sage.
- Nafam (2007) Closing Keynote Address of NAFAM Convention. In: J. Malaysia ed. *National Asset and Facility Management Convention*. PWTC, Kuala Lumpur.
- Ngowi, A. B. and Mselle, P. C. (1998) Community participation in facility management. *Facilities*, 16, (11): pp.314-318.
- O'leary, Z. (2004) *The Essential Guide to Doing Research*. London: SAGE Publications Limited.
- Remenyi, D., et al. (2003) *Doing research in business and management: an introduction to process and methods*. London: SAGE Publications.
- Saunders, M., et al. (2007) *research methods for business students*. 4th ed. Harlow: Pearson Education Ltd.
- Sexton, M. (2004) PhD Workshop: Axiological purposes, ontological cages and epistemological keys. In: *Postgraduate research workshop. November 2004*, University of Salford, UK.
- Sexton, M. (2008) PhD Research: Axiological Purposes, Ontological Cages and Epistemological Keys. In: University of Salford.
- Silverman, D. (1998) Qualitative research: meanings or practices? *Information Systems Journal*, 8, (3): 3-20.
- Sutrisna, M. (2007) Research Methodology: An Introduction. In: University of Salford.
- Tobi, S. U. M. (2003) Inefficiency of Property Management by Local Authority: Case study-Shah Alam City Council. In: *Faculty of Science and Geoinformation*. Kuala Lumpur: Universiti Teknologi Malaysia.
- Tobi, S. U. M. (2006) Outsourcing for Local Authority: Case study-Shah Alam City Council. In: *Faculty of Science and Geoinformation*. Kuala Lumpur: Universiti Teknologi Malaysia.
- Trochim, W. M., 23 July. The Research Methods Knowledge Base, [online]. Available from: <http://www.socialresearchmethods.net/kb> [Accessed: 23 July].
- Wong, T. (2000) Culture-its influence on facilities management in Asia. In: *Proceeding of Facilities Management Asia*. Singapore.
- Yin, R. K. (2003a) *Applications of Case Study Research*. 2nd ed. Thousand Oaks, CA: Sage Publications, Inc.
- Yin, R. K. (2003b) *Case Study Research: Design and Methods*. 3rd ed. London: SAGE Publication

# Translators as Bridge-Builders between Cultures: Fact or Fiction?

**Maria Nader**, *School of Languages, University of Salford*

## Abstract

Translation has been widely regarded as a secondary activity in which words in one language are replaced by others in another language. This misconception has clear implications for the agency of the translator, who is at best viewed as a mute mediator or, at worst, as a robot (Leppihalme 1997: 19). However, nowadays this view is no longer true and both translation and translators are becoming more visible in society. Different studies have recently addressed the issue of the translator's agency, calling into question the widely circulated metaphor of translators as bridge-builders between cultures. This paper contributes to current discussions on the viability of this metaphor and its relevance to the actual role played by translators in everyday translation practice. It will also discuss the idea suggested by Snell-Hornby (1988, 39) that translation is no longer something which occurs between two languages but, more importantly, between two cultures. These ideas are premised on the fact that there is a "vast need for translation in today's world" (Leppihalme 1997: 18-19), and that translators are playing a crucial role as a link between source and target languages. The political, cultural and ideological aspects of language will be studied in the light of the communicative approach set by Newmark (1981), using examples from Arabic, French and Farsi with their English translations.

## 1. Introduction

Communication and dialogue among cultures are quite essential and only possible through understanding the languages of these cultures. For this reason, translation became more needed and translators became the figures who take action towards fulfilling this social need (Leppihalme 1997: 18-19). Nevertheless, translation was regarded as a simple act in which the words of one source text are changed into equivalent target language words through the help of dictionaries. Also, translators were not overly respected in the past, their role was marginalised and it was common to view them as:

"Mechanical devices replacing linguistic codes (equivalents) from one language into another, and the translator's autonomy was always questioned (and is still being questioned) by those who thought of him/her 'as a monkey, with no choice save to make the same grimaces as his master'".

(Leppihalme 1997: 19)

On the other hand, this view changed with the passage of time and people became more aware of the function of translators who create links of understanding between two languages. However, this issue is debatable because it is not always applicable. In other words, different translators interpret differently in that some help make peace, others stimulate conflicts, subdue populations and some fill in language gaps between two nations (Baker 2005: 4). Moreover, the translation process is not as easy as some people might think because of many elements, the most important of which is the translator's familiarity with the target language culture. In fact, the process of translation "can no longer be envisaged as being between two languages but between two cultures involving 'cross-cultural transfer' " (Snell-Hornby 1988: 39-6).

The current paper will concentrate on translation difficulties and on the role of translators and how they deal with such difficulties. The main translation hindrances which will be focused on are those related to culture, politics and ideology. Supporting examples will be provided to show the behaviour of translators who might build bridges or break bridges between two cultures. Also, this paper will show just how interconnected the ideological, cultural and political aspects of translation are. The first issue to be discussed is the relation between culture and translation. Under this title there are two subtitles; the first will explain the cultural

gap between English and Arabic and the second focuses on the limits of translation between English and Farsi.

## 2. Culture and translation

Culture is an important element to be taken into consideration in the process of translation. This is due to the fact that culture is the vessel in which all experiences of a nation are contained such as traditions, habits, thoughts, arts etc. Moreover, "culture is the reflection of the total behavior of a society" (Bennett 1968: 10). Lado (1957: 111) defines it as the "structural systems of patterned behavior" whereas Sapir (1949: 79) argues that "culture is technically used by the ethnologist and culture historians to embody any socially inherited element in the life of man, material and spiritual". On the other hand, different people think in different ways, so different and special mechanisms of translation should be adopted in order to give accessibility to the text translated into the target culture. Moreover, as said before, translators should be aware of the cultural differences between the source and target languages and conscious that they are now translating cultures not simply languages. The following section will concentrate on the cultural hindrances encountered in Arabic/ English translation.

### 2.1. The cultural gap between English and Arabic

There is a big difference between Arabic and English linguistically and this language difference transits to dissimilarities in culture as well. Arabic/ English and English/Arabic translators encounter cross-cultural hindrances that are sometimes difficult to deal with. However, a translator has to be aware of the fact that what is culturally acceptable in one culture is not necessarily acceptable in another one (Bahameed 2008). For example, the idea of having a boyfriend or a girlfriend is common in the Western societies such as the United Kingdom and France but it is totally rejected in the Arab societies. Also, polygamy which is having more than one wife sounds strange to Western societies while it is acceptable in Arab Muslim countries because Islam allows such practice. Differences in culture generate problems in translation especially if the two cultures are widely different from each other. For example, the word 'owl' in an English text symbolizes wisdom but it will be problematic if translated into Arabic, because it is a bad omen and a symbol of bad luck (Ilyas 1989: 124). In such cases, a translator has to seek for appropriate translation procedures to satisfy target language audiences. Such procedures may be additions of other terms to clarify an item, deletion, substitution of the term, or a footnote explaining the different cultural connotations. On the other hand, Ilyas (1989: 124) debates that a translator has to find a good equivalent of the culturally problematic term. For example, to translate an English text containing the word 'wine' into Arabic would be nonsense if it is substituted by, say, 'vodka', because both drinks are not commonly consumed in the Arab countries, plus they are prohibited in Muslim communities.

In fact, some English terms come from the heart of the British society and they do not fit the Arab audience at all. For example, the phrase *one-parent child* is never used in Arabic. Such expression creates a cultural gap between Arabic and English. Homeidi (2004: 17-18) illustrates how a translator can deal with this phrase when it comes in a context such as:

"The **one parent child association** released its latest booklet which includes all the services it offers to its members"

The literal translation of *one parent child* into Arabic will sound opaque as well as failing to convey the correct meaning of the English:

أصدرت جمعية الطفل من أحد الوالدين فقط آخر كتيباتها و الذي يضم الخدمات التي تقدمها لأعضائها

The problem here lies in the fact that there is no formal equivalent of this term in Arabic and in the unacceptability of the notion itself by Arab people. In other words, in the Arab conception, a child with one parent means that he is an orphan who lost one or both parents because of death and that his parents are known through the Civil Service Records. On the other hand, a child who is found without parents is called a 'foundling' not a 'one-parent child' (Homeidi 2004: 17-18). Neither term fits in the Arabic translation; therefore, adding a commentary or a footnote is a necessity to explain the Western concept to the Arab reader. It might read as follows:

"It is possible for a woman in the Western World to have a baby from with (sic) any man she likes, and she is not legally obliged to declare the father's name or nationality etc. or she may not be certain about them. In this case the family which consists of only the mother and the child(ren) is called "one parent family" and the child(ren) is/are called "one parent child(ren)" ".

(Homeidi 2004: 18)

Accordingly, the inability to find cultural equivalents in the target language results in the phenomenon of untranslatability. Bassnett (1998: 79-81) argues that some words and expressions are untranslatable because of the lack of equivalence in the target language. She asserts that translators should not ignore cultural differences but should be aware of them in order to help in crossing the boundaries from one culture to another. Also, for her, it is essential for translators to work with language because language cannot be taken out of culture or culture out of language.

So far, we have discussed the cultural chasm between Arabic and English. The following section will highlight the limits of translation between English and Farsi.

## 2.2. Limits of translation between English and Farsi

One of the characteristics of a translator is to be a mediator, defined by Katan (2004: 171) as someone talented to "understand and recreate culture-bound frames wherever necessary". In fact, translation is a balanced process which depends on many elements such as terminology, source and target texts, and consideration of the target text readers. On the other hand, mediation in translation is defined as "the extent to which translators intervene in the transfer process, feeding their own knowledge and beliefs into their processing of a text" (Hatim and Mason 1997: 147). Consequently, translators or mediators have to present a translated version of the target text through which TT readers can "participate in text production in their own way, seeing connections and meaning instead of stumbling over culture bumps" (Leppihalme 1997: x). In this context, Hatim and Mason (1997: 149) provide an example showing the difficulty of the translator's mediation when it comes to translating from Farsi into English. The example is a political speech by the late Ayatollah Khomeini to the instructors and students of religious seminars in Iran. The text highlights the prominent problems of translating between two languages which are culturally remote from each other. The translator has left most of the expressions as they are in Farsi with very subtle changes in the English text. In fact, there are a lot of sentences that sound unfamiliar to Western readers and express different concepts and notions which reflect the Iranian culture. The following phrases in italics were left unchanged by the translator:

- 1- 'It was through the war that *we broke the back of* both Eastern and Western superpowers'.
- 2- 'It was during the war that we concluded that *we must stand on our feet*'.

(Hatim and Mason 1997: 150)

It should be noted here that some kinds of texts intend to stimulate emotional responses to a certain issue through using emotive or subjective vocabulary (Bahameed 2008). The examples just mentioned came in a context of this kind. The speaker's intention is to arouse his listeners so that they can make a stand towards an issue related to their country. Furthermore, "native speakers of a language have a keen appreciation of the emotive meanings of words" (Bahameed 2008). It is worth mentioning here that the Arabic and the Farsi cultures are close to each other in the sense that Arabic uses the same italicised sentences mentioned above and they are translated easily to both languages, Arabic and Farsi.

In addition, the Farsi text includes a kind of 'over-lexicalisation' which is the use of striking and marked lexical terms e.g. 'money worshippers'; 'parasitical capitalists' and 'world devourers'. These terms were used to identify the external enemies of Iran 'on both political and moral grounds' (Hatim and Mason 1997:151). In the viewpoint of Hatim and Mason the ST did not leave a choice for the translator. He/she was obliged either to leave the text as it is with the risk of being unfamiliar to target readers, or to seek other equivalent terms more congenial to the target audience - but he/she chose the first option (*ibid*). In fact, modifications and changes carried out by the translator sometimes make the original text lose the effect intended by the original writer. In addition, there is a twofold problem in translating such kind

of expressions. The first one is the obvious cultural difference, and the second is the political stand of Iran towards the West, represented by America, which always threatens Iran with war due to the latter's possession of nuclear weapons. Consequently, translation scholars who consider the ideological underpinnings of translation tend to believe that translating itself is a political act (Tahir-Gurcaglar 2003: 113) and the previous examples prove this to be the case.

The next sections of the paper will focus on the idea of translation as a narrative and how this narrative changes in terms of culture, politics and ideology.

### **3. Text modifications for cultural & political purposes:**

In this section, there will be an attempt to show how translators modify texts for cultural and political purposes. Some translators resort to changing certain cultural words; others select what they want to translate and ignore other things. In this way, they either enlarge the cultural gap or make it smaller.

#### *3.1. Translation shifts in subtitles:*

Baker (2006: 1-2) argues that conflict is understood as a state of hostility between groups of people belonging to different races, religions or nation states. Furthermore, in the state of hostility and war, translation and interpreting are considered as "part of the *institution of war* and hence play a major role in the management of conflict - by all parties, from warmongers to peace activists". In addition, in the state of war many meetings, conferences, secret negotiations and seminars are held between conflicting parties so that mediation through translators is required (*ibid*).

However, translators mediate in the text differently either to build or to break bridges between cultures. In the following example, some cultural words are changed for political purposes. It is an English subtitle from Mohammad Bakri's documentary *Jenin Jenin* which was released in 2002 on the occasion of Israel's attack on the Jenin camp in the West Bank. The following analysis of the examples is based on that of Baker (2006: 64-65-66). All the speakers in this documentary were Palestinians of different ages and educational backgrounds. In the following examples, the English subtitlers fail to interpret the Arabic term *shaheed* (martyr) as it was intended by the Arab writer, because its semantic equivalents as chosen in the English version evoke a cluster of Western narratives which give an undesirable interpretation for those who would like to comprehend the documentary.

#### **Example: 1**

**Arabic sentence:** 'لسه بندور شهدا من تحت الأرض'

**Back translation:** We are still pulling **martyrs** from underneath the ground.

**English subtitle:** We are still pulling **victims** out of the rubble.

#### **Example: 2**

**Arabic sentence:** 'خلال دائرة قطرها 30 متر بس وجدنا عشر شهداء'

**Back translation:** Within a circle of no more than 30 meters in circumference, we found ten **martyrs**.

**English subtitle:** In an area of 30 meters we discovered 10 **corpses**.

In both examples, the word 'martyr' was replaced by the words 'victim' and 'corpse' in the English subtitle. Baker (2006: 66) calls this choice of translation a reasonable one because the subtitlers avoided using the word *martyr* since it evokes "associations of Islamic fundamentalism, terrorism and suicide bombing" in addition to being part of the anti-Muslim and anti-Arab storylines spreading in the West.

On the other hand, Arabs believe that whoever dies in a war, whether or not they choose to be involved in it, is called a 'martyr'. Also, the people interviewed in the documentary are against any kind of violence and the meaning of the word *shaheed* (martyr) which they used does not imply any 'militant' sense. Conversely, the West, in general, does not believe in the same idea so words such as 'killed', 'victims' and 'corpses' are the most common in the Western media. Furthermore, Baker (2006: 66) argues that individual words such as 'martyr', 'victim' and 'killed' are not enough to tell the whole story of those people killed, and it would be difficult for the reader to apprehend the whole situation without having a clear idea about the cultural references behind such terms. So, the previous extracts were modified for political and cultural purposes at the same time, but now the following example from the same documentary will show the ability of translators to mediate positively in the text and to build bridges between two different cultures. Baker (2006: 99) quotes a sentence by an old Palestinian man complaining of the global grave silence towards the incursions of Israel in the West Bank. The man says:

**Arabic Version:** "انا عارف والله العظيم، والله العظيم، بيتنا ما صار بيت"

**Back Translation:** 'What can I say, By God, by God, our home is no longer a home'.

**English subtitle:** 'What can I say? Not even Vietnam was as bad as this'.

The literal meaning of the Arabic sentence refers to the loss of lands, dispersal of Palestinians across the world and the endless refugee camps. This meaning lacks intelligibility among the Western audiences and in the USA in particular, since they are far from the region and maybe misled by the media. To clarify this, the subtitlers highlighted the disaster in Jenin Camp using another expression that compared the storyline of Jenin to that of Vietnam War, since the second one is more familiar to the West. The storyline of Vietnam War is more likely to attract the Western reader who is already aware of it. On the other hand, narratives about the suffering of Palestinians are more familiar to Arabs than those about Vietnamese since all Arabs have been influenced by the Palestinian/Israeli war for many years due to geographical factors. The following section will highlight the idea of selective appropriation of some storylines to serve political purposes.

### 3.2. Selective appropriation in the media<sup>1</sup>

Selective appropriation means the translators' manipulation of a text for their own purposes either by deleting or adding information or by changing the whole story of a text. It happens in textual material, literature, interpreting and media. The following example is taken from a web-based translation entitled 'Translation of Funeral' in an Egyptian Paper, *al-Wafd* dated December 26, 2001. It talks about news of the death of Osama bin Laden, the chief of Al-Qa'da Organization in Pakistan. The full version of this example can be found in Baker's *Translation and Conflict* (2006: 119-120). In the original text which was written in Arabic, there is reference to the American bombing of Tora Bora, the region where Bin Laden was assumed to be buried. However, the translator omitted this information in the English version as will be shown in the following sentences:

**Arabic version:** وطبقاً للمذهب الوهابي وأكد المسؤول صعوبة تحديد المكان الذي دُفن فيه (بن لادن) لأن قبره سُوي بالأرض الذي ينتمي إليه. و أكد احتمال طمس مكان الدفن بسبب القصف الجوي المستمر منذ أسبوعين لمنطقة تورا بورا.

<sup>1</sup> Selective appropriation of textual material is realized in patterns of omission and addition designed to suppress, accentuate or elaborate particular aspects of a narrative encoded in the source text or utterance, or aspects of the larger narrative(s) in which it is embedded



**Back translation:** His grave was flattened according to the [tradition of the] Wahabi sect to which he belongs. He confirmed that the burial site might have been obliterated because of the continued aerial bombardment of the Tora Bora region for the past two weeks.

**English Translation:** The official stated that it is difficult to pinpoint the burial location of Bin Laden because according to the Wahhabi tradition no mark is left by the grave.

(Baker, 2006:119-120)

Shavit (1986: 112- 113) opines that a translator is allowed "to manipulate the text in various ways by changing, enlarging, or abridging it or by deleting or adding to it". However, this does not mean deleting or changing the source text as a whole or skewing the meaning completely. It means that a translator can make some modifications to enable TL readers to understand a certain idea or to give them the right information, but not to change the core of the text. In the previous example, the translator deleted important information, namely the bombardment of the Pakistani region in an attempt to play down the American violence. According to Baker (2006), a deletion of this kind is called selective appropriation, because the translators chose what suited their purpose neglecting other important issues that should have been transferred as well. In fact, "public narratives promoted by powerful institutions such as the state or media not only highlight those elements they selectively appropriate, but also force them on our consciousness through repeated exposure" (Baker 2005: 9). In other words, the media plays a prominent role in changing world opinion by the repeated portrayal of a party as being persecuted while it is the one which practices persecution against others.

The second explanatory example is based on a paper by Quentel (2006). In his paper, Quentel analyzes the translation of G.W. Bush's 2003 State of the Union Address related to the war against Iraq. *Le Monde*, a well-known French newspaper, translated part of the speech. France at that time opposed war in Iraq. The translated version in *Le Monde* shows how French journalists were biased and selective in translating Bush's speech. There was a play with phonetics in order to produce a specific effect. According to Quentel (2006: 3), a translator of political speeches has to be sensitive to norms such as the social, ideological and historical contexts of a speech rather than to concentrate on the degree of semantic accuracy which is usually low in such kind of texts. Indeed, the anti-American or more specifically the anti-Bush feeling was deep-seated in French society; and the political situation between France and the USA was tense at the time Bush delivered the speech. However, *Le Monde* was not completely affected by this anti-American trend and tried to be as neutral as possible, but the translation was not one hundred percent free of anti-Bush connotations. In fact, there were two French translations of the English original speech, an official one published by the US Department of State and the other one published in *Le Monde*. Though there is not much difference in the content of both versions, their results nevertheless sound different. Some of the differences appear in examples such as:

'To date, we've arrested or otherwise **dealt with** many **key commanders** of Al Qaeda'

In this sentence, the phrase 'dealt with' was translated by *Le Monde* as "réglié le sort" opposite to the official translation: "nous les avons neutralisés". The first one expresses the idea that Bush acts like a cowboy in a Western movie using strength only to dominate rather than a good politician acting wisely.

Also, the phrase 'key commanders' is translated by *Le Monde* as only 'chefs' while it is 'grands chefs' in the official translation. Quentel (2006: 6) explains this change as a way of undermining the work of the American campaign against Al-Qaeda.

So far, the *Le Monde* translation is a reflection of the anti-Bush or the anti-American feeling which pervaded in France at the time. In this context, Leppihalme (1997: 3) confirms that "texts occur in a given situation in a given culture in the world, and each has a specific function and an audience of its own". Moreover, Tymoczko (2003: 182-83) argues that "the ideology of translation resides not simply in the text translated, but in the voicing and stance of the translator, and in its relevance to the receiving audience". In other words, some texts serve the interests of their writers or readers as will be shown in the following example:

"The budget I send you will propose almost \$6 billions to quickly make available effective vaccines"

This sentence said by Bush was translated similarly in the official French translation, but *Le Monde* provided a different interpretation:

"Le budget compta 6 milliards de dollars pour obtenir rapidement des vaccins"

According to Quentel (2006: 7), the reader of the second sentence by *Le Monde* will think that Bush has decided alone that the budget will be \$6 billion without getting back to the Congress, a high democratic entity in the USA, which can vote for or against the decisions taken by the president. So, in this case Bush will be viewed as disrespectful to democratic institutions. In fact, this translation meets the expectations of the French people and reflects the general impression about Bush in France. In this way, translation can be described as a two-edged sword which either builds or destroys so translators are supposed to use it wisely.

#### **4. Translation shifts on cultural/ideological backgrounds**

##### *4.1. The relationship between culture and ideology*

Khol (1984: 17) defines culture as "an integrated system of learned behavior patterns that are characteristic of the members of any given society". On the other hand, ideology is defined as "a belief or a set of ideas, especially the political beliefs on which people, parties, or countries base their actions" (Collins Cobuild s.v). Calzada- Pérez (2003: 6) argues that "Cultures are often regarded as traditions, parts, roots or knowledge, in short, heritages". Moreover, ideology is defined as a framework that is "assumed to specifically organize and monitor one form of socially shared mental representation, in other words, the organized evaluative beliefs—traditionally called 'attitudes'—shared by social groups" (Van Dijk 1996: 7). Yet, "Cultural differences tend to take over from ideological differences and become practically indistinguishable from them, especially when there is a clash of civilizations" (Baker 2006: 41-42). All these definitions denote that ideological and cultural aspects are almost the same and they probably cause the same problems when it comes to translation. Sometimes ideology is viewed in a more positive sense when it becomes a means of helping and supporting a particular social group rather than destroying or discouraging it (Calzada-Pérez 2003: 5). Also, cross-cultural and ideological tensions are an old issue found in all societies (*ibid*). However, ideology was always associated with 'negative-political-connotations' while culture was traditionally connected with positive features. Álvarez and Vidal (1996: 5) argue that behind every one of the translator's selections, such as what to add, what to leave out, which words to choose and how to place them, "there is a voluntary act that reveals his history and the socio-political milieu that surrounds him; in other words, his own culture [and ideology]". In addition, since being critical of one's own culture can be seen by some as 'risky' and 'inappropriate', it is also 'politically incorrect' to criticize other cultures openly (Calzada-Pérez 2003: 6). In other words, to criticize my own culture is usually undesirable and consequently criticizing and assessing another culture publicly will evoke bad feelings towards me (as a critic). So, there are three things a translator has to consider when interpreting a text: the cultural, political, and the ideological parameters of the target language society. However, some translation scholars think it unnecessary to modify the text according to the mood of the TL audience, which should be introduced to new things coming from the SL culture. This idea will be discussed in the following section.

##### *4.2. Domestication (ethnocentric) vs. foreignization (ethnodeviant) of translation*

Venuti (1995) has distinguished between two concepts: 'domesticating' translation and 'foreignizing' translation. He has written about how Anglo-American translation has deprived the source text of its flavour through putting it in a mould that is so familiar to the target culture. In this way, the modified text becomes unchallenging, because the cultural values of the source text will be also modified, and consequently lose its original flavour (Hatim and Mason 1997: 145). A translator has to mediate by only familiarizing the opaque terms that appear in the source text without changing the whole meaning of the original text. In this way, a line of connection is created between both cultures without distorting or underestimating any

of the languages concerned. So, despite the dissimilarities between any two civilizations, a translator is always invited to fill in the ideological chasm that separates both of them.

A good example of changing parts or lexical terms in translation is the story of Snow-White by the Brothers Grimm. There would be some modifications in parts of the story if translated into Arabic. The main thing that would be deleted from the story is the word 'wine' in the scene in which Snow White first came to the dwarfs' house and drank a bit of wine. An Arab translator would either delete the word or substitute it with an alternative like 'orange juice'. That is because the concept of drinking wine is forbidden for Muslims, who form the majority in the Arab Countries. In this case, if the alternative 'orange juice' is selected, then it is the translator's attempt to domesticate the term 'wine' by substituting it with something familiar to the target audience, as well as to avoid any clash with the culture itself. Also, the scene where the dwarfs wash Snow White's body would not be translated because it would sound very obscene to Arab children who are not used to seeing such intimate scenes. A translator might say that the dwarfs washed her hair, and that would be enough for children. In this context, Baker (2006: 64) asserts that "Translators and interpreters at times also avoid the use of a direct semantic equivalent of an item in the source text or utterance when that equivalent is or has become embedded in a different and potentially negative set of narratives in the target culture". This applies to the term 'wine' which is usually connected with negative narratives in the Arab world such as losing mind and motivating lust if drunk heavily.

## 5. Conclusion

To sum up, the three translation hindrances dealt with in this paper, namely ideology, politics and culture, are all inseparable and have to be taken into consideration when it comes to translating different text genres. This does not contradict with other important elements in the process of translation such as audience education, language of the text, age of readers and other factors which come out of the scope of this paper. As we have seen in the paper, some translations mislead readers as in the example about bombarding Tora Bora; and other translations try to influence readers through using their existing familiarity with key events, as in the example where the attack on Jenin camp was compared to the Vietnam War. However, translation is subjective in that translators may change certain narratives for certain purposes. So, the answer to the main question of this paper about whether or not translators are able to create bridges between cultures is actually debatable. Nevertheless, translators should not be outside or in between cultures but at the heart of interaction and the storylines they translate (Baker 2005: 4-13). Therefore, the role of translators and interpreters should not be partial any more. They have to be culturally educated and responsible of what they translate; otherwise, they will lose credibility with readers and cause misunderstandings that may lead to the clash of civilizations.

## References

- Bahameed, A. S. (2008) 'Hindrances in Arabic-English Intercultural Translation', *Translation Journal*, vol. 12, No. 1, January.
- Bennett, W. A. (1968) *Aspects of Language and Language Teaching*. Cambridge: Cambridge University Press.
- Baker, M. (2005) 'Narratives in and of Translation' *SKASE Journal of Translation and Interpretation* 1(1): 4-13.
- Baker, M. (2006) *Translation and Conflict: A Narrative Account*. London & New York: Routledge.
- Bassnett, S. (1998) 'Translation across Cultures', in *Language at Work*, British Studies in Applied Linguistics, pp: 72-85.
- Calzada-Pérez, M. (2003) 'Introduction', in Calzada-Pérez, M. (ed.) (2003) *Apropos of Ideology*. Manchester: St. Jerome Publishing. P: x-z.
- Hatim, B. & Mason, I. (1997) *The Translator as Communicator*. London: Routledge.
- Homeidi, M.A. (2004) 'Arabic translation across cultures'. *Babel*, Volume 50, Number 1, pp. 13-27(15): John Benjamins Publishing Company.
- Ilyas, A. (1989) *Theories of Translation*. Mowasil: University of Mowasil.

- Katan, D. (2004) *Translating Cultures: An Introduction for Translators, Interpreters and Mediators*. Manchester: St. Jerome.
- Lado, R. (1957) *Linguistics across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor: University of Michigan Press.
- Leppihalme, R. (1997) *Culture Bumps: An Empirical Approach to the Translation of Allusions*. Clevedon: Multilingual Matters.
- Quentel, G. (2006) "The Translation of a Crucial Political Speech: G.W.Bush's State of the Union Address 2003 in Le Monde", University of Warwick, CSGR Working Paper, n.204
- Sapir, E. (1949) *Culture, Language and Personality*. California: University of California.
- Shavit, Z. (1986) *Poetics of Children's Literature*. Athens: University of Georgia Press.
- Snell-Hornby, M. (1988) 'Translation Studies: An Integrated Approach', in Katan, D. (2004) *Translating Cultures: An Introduction for Translators, Interpreters and Mediators*. Manchester: St. Jerome. P: 170.
- Tahir-Gürçağlar, Ş. (2003). 'The Translation Bureau Revisited: Translation as Symbol', in Calzada-Pérez, M. (ed.) *Apropos of ideology*. Manchester: St. Jerome. (P: 113-130).
- Tymoczko, M. (2003) *Ideology and the Position of the Translator - In What Sense is a Translator 'In Between'?* In Calzada Pérez, Maria (ed.) *Apropos of Ideology. Translation Studies on Ideology - Ideologies in Translation Studies*. Manchester: St. Jerome Publishing, pp. 181-201.
- Van Dijk, T.A. (1996). *Discourse, Opinions and Ideologies*. In Schäffner C. & Kelly-Holmes H. (Ed.) *Discourse and ideologies* Philadelphia: Multilingual Matters. pp. 7-37.
- Venuti, L. (1995) *The Translator's Invisibility. A History of Translation*. London: Routledge.

# Identifying and classifying stakeholders of post-disaster housing reconstruction projects in Sri Lanka

**Nuwani Siriwardena, Richard Haigh and Bingunath Ingirige**, School of the Built Environment, University of Salford

## Abstract

A construction project is a temporary, one time only and short term undertaking that creates the built environment. Stakeholders of a construction project are any identifiable group or individual who can affect, or is affected by, the achievement of a project's objectives. Hence, stakeholders vary depending on the nature of the project undertaken. A disaster frequently leads to the destruction of buildings and infrastructure, including houses, thereby disrupting livelihoods and affecting the personal, social and economic lives of the victims, and hindering development of the country. The 2004 Boxing Day Tsunami caused massive destruction in Sri Lanka and made 98,000 people homeless. Providing houses, which was a vital part of re-instating the livelihoods of people displaced by the Tsunami, became a challenging task and gave birth to a large number of post-disaster housing reconstruction projects. A wide array of stakeholders came together in such projects to execute the final outcome. In the aftermath of early reconstruction efforts, some stakeholders expressed dissatisfaction. Areas of criticism included time, cost, quality, coordination with infrastructure and linkage to livelihoods. There is now growing recognition of a need to more effectively identify and classify stakeholders of post-disaster housing reconstruction projects so that expectations can be more successfully managed. This paper presents the interim results of a study to identify and classify the stakeholders of post-disaster housing reconstruction projects in Sri Lanka, based on an extensive review of the literature and pilot interviews with stakeholders involved in housing reconstruction projects in Sri Lanka. The results identified seven stakeholder groups: the client; contractor; consultant; donor; NGO's; INGO's; CBO's; and, the beneficiary of post-disaster housing reconstruction in Sri Lanka. Primary-secondary, internal-external, voluntary-non voluntary, and social-non social were identified as the four main classifications of stakeholders that emerged from the literature. Interviews were used to categorise the identified stakeholders under a suitable classification. Stakeholder expectation gaps of post disaster housing reconstruction in Sri Lanka were identified under three main themes: problems relating to design and the quality of the house; problems of relocation; and, institutional problems.

## 1. Introduction

In recent years disasters have been affecting increasing numbers of people throughout the world, while budgets for emergency and humanitarian aid have swelled (Centre for Research on the Epidemiology Disasters, 2007). Over the past century, disasters have grown increasingly complex in their origins, as well as their impacts on the environment, economies and human beings (Peek and Sutton, 2003). In 2007, 414 natural disasters were recorded (CRED, 2008). More than 16 thousand people died and over 234 million were affected, causing almost US\$ 75 billion in economic damages (CRED, 2008). Asia was the worst hit, where 4,234 were reported killed in Bangladesh by cyclone Sidr, and more than 3,000 fatalities were recorded from severe floods in Bangladesh, India, North Korea and China (UN/ISDR, 2008).

One of the deadliest disasters in recent history was the earthquake that struck off the Coast of Sumatra Island in December 2004, which had a magnitude of 9.0 on the Richter scale (GOSL, 2005). This rupture caused a tsunami, a series of water waves created when a large volume of a body of water is rapidly displaced. The year 2004 saw the highest number of people reported killed (Hoyois et al., 2007). This was inclusive of 157,564 reported deaths, 26,763 reported missing and 1,075,350 reported displaced by the 2004 Indian Ocean (Boxing Day) Tsunami (UN/OCHA, 2005). Indonesia, India, Sri Lanka and Thailand were the hardest hit.

The tsunami waves that struck nearly 2/3 of the Sri Lankan coastal belt on 24<sup>th</sup> December 2004 resulted in 35,320 confirmed deaths and 6,300 missing people. The economic loss has been estimated at US\$ 2.2 billion, disregarding the cost of social disorder. Nearly 516,100

were displaced, of whom 33% were below the poverty line. 98,000 dwelling houses and 75% of the fishing fleet, which is the main livelihood of the coastal population, were damaged. More than 500 tourism related businesses - the main source of income in this area - were damaged. Further, 23,449 acres of agricultural lands were salinated; 200 Educational Institutes were damaged and 450 others were used as temporary camps for displaced people. The world's worst train tragedy, killing 1700 passengers, also resulted from the Tsunami. Within this context, it necessitated operations to rehabilitate, recover and reconstruct the affected community. Moreover, the Department of Census and Statistics (DSC) in Sri Lanka confirmed that thirteen districts of the five provinces on the coastal belt were affected. Table 1 provides a summary of the damage caused to housing.

**Table 1** Destruction to houses in different provinces

Province and Districts	Number of houses completely damaged	Number of houses partially damaged-cannot be used	Number of houses partially damaged-can be used	Total number of houses affected
Southern Province-Matara, Galle, Hambantota	7,222	2,116	10,971	20,309
Western Province-Kalutara, Gampaha, Colombo	5,107	1,285	6,548	12,940
Eastern Province-Ampara, Batticaloa, Trincomalee	19,447	5,637	18,632	43,746
Nothern Province-Jaffna, Mullatiyu, Killinichchii	8,085	1,026	2,383	11,494
North Western Province-Puttalam	12	16	27	55
Total	39,903	10,080	38,561	88,544

Source: Department of Census and Statistics Sri Lanka

The task of reconstruction after a major event can be an onerous challenge to a developing country like Sri Lanka, which requires the deliberate and coordinated efforts of all stakeholders for effective and efficient recovery of the affected community. Re-housing formed one of the greatest challenges for Sri Lanka. According to the UN, 516,159 people were made homeless and 98,000 homes were destroyed or suffered major damage. A year later, 4,299 new houses had been built and another 25,000 were planned, of which 10,707 were under construction. A start had yet to be made on planning the remaining 68,000 dwellings (GoSL, 2005). Similarly, Oxfam reported that 12 months after the Tsunami, around 20% of the people made homeless were not yet in satisfactory permanent accommodation (Oxfam International, 2005, p6). A year and a half after the tsunami, the situation appeared little better, with most people who had lost their homes still without a permanent roof over their heads and many still facing an uncertain future. This is despite the record public and donor response resulting in a quick and largely successful immediate human relief effort. In 2006, approximately 200,000 people were still displaced (ICRC, 2006). Many stakeholder groups expressed dissatisfaction in relation to features of the housing, relocation and other institutional problems.

The broader aim of this study is to explore and investigate how to identify, classify and manage stakeholders, and their expectations, in order to deliver effective post-disaster housing reconstruction in Sri Lanka. This paper presents the interim findings of the study, based on an initial set of interviews and a review of the literature. Methods to identify and classify stakeholders are considered in the context of post-disaster housing reconstruction. The

study goes on to explore stakeholder expectation gaps in post-disaster housing reconstruction in Sri Lanka.

## **2. Literature review**

### *2.1 Identifying and classifying stakeholders from a construction point of view*

A construction project is a temporary, one time only and short term undertaking that creates the built environment. Walker et al. (2008), defined stakeholders of a construction project as individuals or groups who have interest or some aspects of rights or ownership in the project, and can contribute to or may be impacted by, either the work or outcomes of the project. Representatives of different and sometimes discrepant interests are regarded as stakeholders of a construction project (Olander and Landin, 2005). This implies that 'stakeholder' is a collective noun embracing a wider group of people. Newcombe (2003) has accommodated in his definition of stakeholders 'groups or individuals who have a stake in, or expectation of, the project's performance and include clients, project managers, designers, subcontractors, suppliers, funding bodies, users and the community at large'. Stakeholders of a construction project are affected by or affect the development of the project. Hence, capturing their input is a key component of the project development process. It is noted that there are common parties coming under the spectrum of construction and post-disaster housing reconstruction project contexts.

Winch (2002) identified internal stakeholders of a construction project as the parties who have a legal contract with the client and external stakeholders as a group who possess a direct interest in the project. CIOB categorise construction related stakeholders into three groups, namely business partners composed of employees, suppliers, distributors, service providers, unions; authorisers which include government (legislation and guidance), regulatory agencies (e.g. HSE), shareholders / investors, professional bodies (e.g. CIOB), boards of directors (in large companies); and external influencers consisting of consumers / clients (current/end users of the building), public (communities affected by construction), media (e.g. journalists), special interest or pressure groups (e.g. environmental/social welfare groups). This classification identifies three groups with a special concern in the construction industry. Walker et al. (2008) divides stakeholders into four groups: upstream supply chain partners; downstream supply chain partners; external stakeholders; and, project stakeholder group. According to Walker et al. (2008), upstream stakeholders include the paying customer and end users of the product/service; downstream supply chain partners consist of suppliers and subcontractors. In contrast, external stakeholders include general community and independent concerned individuals or groups who may be invisible and feel that they will be affected by the project and its outcomes and engage with the project team in delivering the ultimate project benefit, but whose cooperation and support is vital for project success, and also knowledge network members who interact with the project delivery team in a variety of ways. Lastly, the project stakeholder group comprises of the project sponsor or champion as well as the project delivery team. Therefore, it can be deduced that stakeholders of a construction project are any identifiable group or individual who can affect, or is affected by, the achievement of a project's objectives.

Differing classifications of stakeholders have emerged due to many reasons. Mitchell et al. (1997), highlighted that different classes of stakeholders can be identified by possession of attributes, such as: the stakeholder's power to influence the firm; the legitimacy of the stakeholder's relationship with the organisation; and, the urgency of the stakeholder's claim on the organisation. Further, Corporate Social Responsibility (CSR) and Corporate Governance (CG) have opened the avenues to identify different stakeholders and their influence over corporate decision making (Bakker and Hond, 2008). Nevertheless, an organisation has a role and responsibility to deliver the expectations of different categories of stakeholders (Carroll, 1999). Thus, stakeholders are conceptualised as having direct relationships with one another and relationships emerge depending on the context and necessity (Carroll, 1999; Bunn et.al., 2002; Andriof and Waddock, 2002; Winch, 2002). It can be concluded that the categorising of stakeholders is socially constructed.

Varying understandings of stakeholders have given rise to differentiating among stakeholders depending on the power, legitimacy and urgency to the project (Podnar and Jancic, 2006). Stakeholder power represents to what degree a stakeholder can cause a change to the project. The sources of power could be positional, political or personal. Legitimacy indicates the extent to which the stakeholder is involved in the project. There may or may not be a formal contractual agreement. Urgency reflects to what degree the stakeholder is important, necessary or desirable. This could reinforce the positive aspects of the relationship or lead to mitigate negative influences. Changes in stakeholder relations occur due to predictable, unpredictable, internal and external factors. This paper will focus upon stakeholders that emerge following a disaster, which could be predictable or unpredictable, and where internal or external factors may influence the reconstruction projects.

## *2.2 Stakeholders in post-disaster housing reconstruction in Sri Lanka*

The precise identity of the stakeholders will evolve depending on the context and the situation (Carroll, 1999; Bunn et al., 2002; Andriof and Waddock, 2002). Thus, stakeholders of post-disaster housing reconstruction and their expectations are likely to be different to a normal construction project. For example, for housing in the post-disaster situation, there are some added challenges: the scene is generally very chaotic and resources are in scarce supply, with simultaneous projects being launched by numerous local and international organisations for housing and infrastructure repairs, for creation of livelihoods, and for a range of other social programmes; projects must be completed as quickly as possible to foster recovery and to satisfy donors who want to see results; and, the post-disaster period is generally seen as a good opportunity to engage in activities that will increase the level of development and reduce vulnerability to future disasters, implying that projects must be implemented with sustainability in mind (Davidson et al., 2007). Similarly, Sundnes and Birnbaum (2002) describe disasters as a mismatch between resources and tasks, whether caused by natural or avoidable man-made (non natural, human induced) events. The impact of a disaster, causing death, injury, loss of property, economic losses etc. is felt when it overwhelms the community's ability to cope (Twigg, 2001; Christian Aid, 2003). Limited resources hinder effective disaster management. Thus, lack of resources typically prompts international assistance flow to the countries affected (Hopkins, 2005; GFDR, 2005).

International Non-Governmental Organisations (INGO's) such as Oxfam and Care International, provided medical assistance. Funds flowed from donor agencies such as the Asian Development Bank, World Bank, International Committee of Red Cross and UN Habitat. The national government made the policy decisions related to the reconstruction and the local government executed the need assessment plans, in coordination with the donor agencies, Non Governmental Organisations (NGO's), Community Based Organisations CBO's, humanitarian organisations, professional organisations and the national government.

However, in such circumstances, a country like Sri Lanka typically lacks strategies for delivering effective disaster reconstruction and managing the associated resources. There is however growing recognition that the construction industry has a responsibility to assist communities to anticipate, assess, prevent, prepare, respond and recover from disasters (Haigh et al., 2006) by contributing in developing, implementing and operationalising the appropriate strategies. In order to do this, stakeholders of construction projects composed of professionals such as engineers, quantity surveyors, architects, project managers, clients and contractors, have to work hand in hand. Moving beyond a disaster connects other stakeholders such as the NGOs, INGO's, donor agencies, CBO's and the beneficiaries with construction project stakeholders. It is also important to identify a suitable classification for the wide array of above mentioned stakeholders.

Within the literature, four main classifications of stakeholder are identified: primary-secondary, internal-external, voluntary-non-voluntary, social-non-social. Primary stakeholders are parties who derive a benefit directly out of the project. Secondary stakeholders are indirectly affected parties who act as intermediaries. Internal stakeholders of a construction project are the parties who have a legal contract with the client. External stakeholders are a group who possess a direct interest with the project. Voluntary stakeholders have a contractual relationship with the project. Non-voluntary stakeholders do not have a contractual agreement. Social stakeholders are actively involved in the decision making process. Non-



social stakeholders are unlikely to have a direct power to get involved in the decision making process.

### *2.3 Stakeholder expectation gaps in post-disaster housing reconstruction*

It is widely accepted that the impact of disasters is greater in developing countries, where the reconstruction process takes a prolonged period (UNDP, 2005; World Bank, 2006). Damage and destruction to housing, infrastructure and other facilities restricts the development of the economy while increasing the likelihood of epidemics. Environmental degradation is also common. It takes a prolonged period to reinstate the personal, economic, and social lives of the affected due to a lack of financial and intellectual resources (Keraminiyage et al., 2006; Pardasani, 2006; WHO, 2006; Perry, 2007). Recovery planning, information management, risk minimisation, participatory approaches, regulation, and awareness-raising are ill developed in these countries, including Sri Lanka (Jigyasu, 2002; 2004, Oloruntoba, 2005; Billa et al., 2006; Kurita et al., 2006; Mulligan and Shaw, 2007; Boano, 2009). Effective disaster management in Sri Lanka is also threatened by institutional constraints, gaps in communication, a lack of access to professional skills and knowledge to support local effort, and a failure in management and planning (RICS, 2006). Further, the civil conflict which continued for more than three decades has slowed down the process of post-disaster housing reconstruction.

Moreover, the quality of the resultant housing constructed has also been questioned. In Sri Lanka, the President celebrated the fact that 50,000 transitional shelters had been put up in the first six months but, with the start of the monsoon season, many of these homes were already in need of upgrading (Oxfam International, 2005, p.5). Similarly, an evaluation of CARE, Oxfam and World Vision's response to the tsunami in Sri Lanka by Bhattacharjee et al. (2005) found a series of flaws in the design and implementation of housing programmes (ProVention/ALNAP, 2005a, p.6).

Nissanka et al. (2008) and Ratnayake and Rameezdeen (2008) identify stakeholder expectations in post-disaster housing reconstruction as an under-researched domain where further in-depth studies are required. Only a few studies have been carried out to explore the expectations of stakeholders in a disaster in Sri Lanka. For example, evidence collected from the victims of the droughts in 2001 shows that they expected the authorities and public to assist them, but the nature of the support expected has not been adequately analysed or understood (Ariyabahu and Hulanagmuwa, 2002). Further, local expectations of the type of shelter to be provided through the reconstruction program differed from what the intervening agency was prepared to provide. For example, permanent houses that were built by certain NGOs in aftermath of the Tsunami 2004 failed to consider the needs of the beneficiaries and the quality was inadequate (GOSL, 2005). Hayles (2008) confirms that hasty decisions made without appropriate local consultation inevitably create inappropriate and unwanted results, such as housing which cannot be used by local communities, and housing that does not meet local standards for safe building.

A sustainable livelihood could mitigate the mental trauma and unrest among victims as it would ensure a regular source of income. The process of livelihood recovery is a slow, highly complex undertaking that frequently involve factors outside the control (and competence) of international humanitarian relief agencies (TEC<sub>a</sub>, 2006, p.16). Thus, relocation of tsunami victims who were within the no built area (costal buffer zone), were victimised twice as uncertainty of livelihood developed within them (Perry, 2007, Lyons, 2009). Many were engaged in fishing, coir industry and other small entrepreneurship. Moving away from the coast created unrest among the bread-winners of the affected families as well as the other members. The question of what would be their source of income afterwards, cropped up in their minds. Re-building communities and livelihoods is more complex and takes longer than building houses or distributing goods (Telford et al., 2006).

Compatibility of socio-cultural and religious norms would lead to fewer or no conflict situations. Literature on the relocation of housing reported many incompatibilities among communities affected as well as the neighbourhoods (Lyons, 2009). Thus, the differences of cultural and social patterns / behaviour lead to a fresh set of problems in the neighbourhood (Lyons, 2009). In particular, the people from the coastal area are believers of Christianity whose attitudes, values, beliefs and behaviour, are different to others who are living outside the coastal area. Thus, the mismatch between these socio, cultural and religious norms may

create the space for a conflict with the people who reside inland. Hence, the neighbourhoods of certain inland areas have publicly expressed their dissatisfaction with regard to tsunami victims moving to their territory (Fazlulhaq, 2008). Moreover, the eastern part of the country contained the two distinct ethnic communities of Muslims and Tamils, creating a complex and mixed geography. Authorities had to consider the concentration of these ethnic groups when resettlement decisions were made.

It was highlighted in Provention (2005) that, rebuilding should be in situ wherever possible, as experience has shown that affected people usually want to stay close to their original homestead. Moreover, local customs including the needs of extended families, the location of the kitchen, the building of verandas, privacy and weatherproofing should also be taken into account when houses are put up in areas different from their original positions. These issues are directly linked to the selection of land and the design of the house. Thus, construction professionals have a role to play in this regard, coordinating with other stakeholders.

Infrastructure and other services form a major element of post-disaster housing reconstruction projects. For the purposes of these projects, infrastructure was composed of highways, railways, bridges, water and sanitation, electricity and services including schools (pre-school, primary and secondary), healthcare facilities, postal service and transport services. Donors were also responsible for provision of infrastructure including water, power and sewerage services (TAFREN, 2005) which they failed to provide (GoSL, 2006). For example a research study revealed that over 60% of all the newly relocated occupants surveyed found the amenities, including water, worse than before the tsunami (Weerakoon et al., 2007). Moreover, at a tsunami shelter in Lunawa, a suburb of Colombo, 93 families share a single toilet. In the east coast town of Kalmunai in Ampara District, the Jiffrey welfare centre has only one water tank for the whole camp. A study carried out by International Labour Organisation in 2007, highlighted that about 1,250 houses built by donors remained vacant because of lack of security, infrastructure and limited income generating opportunities.

The root cause behind all institutional problems is the undue political influence (Oloruntoba, 2005; Perry, 2007; Weerakoon et al., 2007). All strategic level decisions were made by the government. The Government decision making process itself is highly bureaucratic and dominated by red tape in Sri Lanka. The tendency of incoming governments to establish new ministries on top of existing structures has led to the creation of new government agencies with overlapping policy responsibilities. Politicised decision making processes have given rise to many of institutional problems including the predictability of the buffer zone, tsunami housing policy etc. Further, the Tsunami led to a sudden availability of an enormous amount of funds, where spending plans were absent, and resulted in inequitable distribution of funds. These funds exacerbated mistrust, facilitated corruption, and led to the resumption of large scale conflict because of the uneven geographical spread of these funds (Rex, 2006; Weerakoon et al., 2007). These and other similar reports suggest that in Sri Lanka, as in other developing countries affected by disasters, much of the post-disaster reconstruction effort has failed to adequately meet the needs of stakeholders in the affected communities.

It has been shown that there are families still living in temporary shelter since 2004, in certain western parts of Sri Lanka. They have not even received any assistance to obtain a house for the house they lost. Further, they are threatened with evacuation from the places where they are currently residing (Relief web, 2008; Fazlulha, 2008). It has also been reported that basic facilities and services such as water, electricity and sanitation have been discontinued from September, 2008 (Relief Web, 2008).

There are families who are still in the process of obtaining permanent housing (Fazlulhaq, 2008). Though they have received money from the funding body, they have been deceived by the sellers of land in cases where the municipality will not permit the construction of a house on that land.

Oxfam (2005) found that post-tsunami aid had predominately gone to the businesses and landowners, while poor people had been neglected in Sri Lanka. Similarly, Perry (2007) reported one community representative in Sri Lanka who expressed anger at the delays in the provision of better housing by several aid agencies that "did not care for the people and built very nice big offices for themselves instead of his village." Further, the Tsunami led to a sudden availability of enormous amount of funds, where spending plans were absent, and this resulted in inequitable distribution of funds. These funds exacerbated mistrust, facilitated corruption, and led to the resumption of large scale conflict because of their uneven geographical spread (Rex, 2006; Weerakoon et al., 2007). These and other similar reports

suggest that in Sri Lanka, as in other developing countries affected by disasters, much of the post-disaster reconstruction effort has failed to adequately meet the needs of stakeholders in the affected communities. Thus, ineffective and inefficient project management and lack of resources have become contributing factors to sluggish reconstruction efforts.

From the literature there is an apparent dissatisfaction from stakeholders, and therefore a need to further understand the gaps in stakeholder expectations and their underlying causes. Therefore, in the journey of exploring and investigating how to identify, classify and manage stakeholders and their expectations in order to deliver effective post-disaster housing reconstruction in Sri Lanka, above expectation gaps were identified in literature. The results of the preliminary interviews are presented in the section on 'Findings', following 'Research Methods'.

### 3. Research Method

An extensive survey of the literature and pilot interviews were the research methods adopted for this study. The literature review section (2.1) stipulates the importance of identifying and classifying stakeholders from the point of view of construction, which laid the foundation for identifying the stakeholders of post-disaster housing reconstruction projects and recognising a suitable classification. Further, stakeholder expectation gaps pertaining to post-disaster housing reconstruction were identified from the literature.

Fifteen open ended questionnaires were distributed among the personnel who got involved in post-disaster housing in Sri Lanka, representing the government, private sector, NGO's, donor agencies and the community, out of which seven responded and consented to a brief interview. Seven pilot interviews provided rich data on the real life experience of the interviewees. Data was analysed based on the interpretations. Findings are discussed in detail in the following section.

### 4. Findings

Seven pilot interviews enabled testing of the data collected and analysing strategies. Further, these formed the basis for the next phase of the study. The findings of the preliminary interviews are discussed under three main headings:

- A. Identification of stakeholders and their engagements
- B. Classification of stakeholders in the context of post-disaster housing reconstruction
- C. Stakeholder expectations and gaps identified in post-disaster housing reconstructions in Sri Lanka categorised under three main headings; time, cost, quality and design of the house, problems of relocation and institutional problems (which are explained in detail below).

#### A. Identification of stakeholders and their engagements

Providing houses, which was a vital part of re-instating the livelihoods of people displaced by the Tsunami, became a challenging task and gave birth to a large number of post-disaster housing reconstruction projects. A wide array of stakeholders such as the government, the private sector, NGO's, INGO's, CBO's and humanitarian organisations came together in such projects to execute the final outcome. The type of engagement is summarised in the Table 2.

**Table 2** Different stakeholder groups identified

Stakeholder group	Type of engagement
1.Client-Government	Client can be public or private. In most cases it is the government. Disaster is a social phenomenon. In the context of a disaster, the government takes the lead in terms of formulating and maintaining regulations, policies and monitoring the adherence to these. Setting the

	standards relating to the delivery of post-disaster housing reconstruction projects is also a vital role played by the government. E.g. Buffer zone imposed by the government after Tsunami. The main difference between a normal construction project and a community project is that the client and the beneficiary is the same in a normal construction project and in a post-disaster housing reconstruction project the main initiator is the government and benefit accrues to the community affected.
2.Consultant	Provides the consultancy advice for the project on designing, evaluating the cost, technical issues/advice (engineering advice, electrical, civil etc.)
3.Contractor/sub contractors	Engage in actual construction according to the designs, specifications, contract documents communicated by the relevant parties
4.Funding body/donor	E.g. UN, ADB, ICRC. Address humanitarian issues while providing the necessary funds to the community project. Ensures that the funds are utilised for the purpose. E.g. if a pre- condition is imposed to spend the money on community development, the donor has to make sure that the funds are used for this particular activity.
5. NGO's (Local)	Acted as the mediator between the funding body and the government
6. INGO's	<p>Developed and incorporated guidelines to strengthen local capacity and ensure accountability to vulnerable populations receiving humanitarian and development assistance</p> <p>Assisted in constructing tens of thousands of temporary shelters and permanent homes and creating livelihood programs</p> <p>Supported microcredit and other programs designed to empower local communities.</p> <p>Provision of expert advice and other resources to reinstate the affected community. INGO assistance included provision medical assistance, counselling for the people suffered from mental trauma, etc.</p> <p>Educated the public and the media on the components and nature of effective and sustainable disaster preparedness and response.</p>
7.Beneficiary/End user	The most important stakeholder. Since they are the beneficiaries, their engagement should be to communicate their needs/ requirements to the relevant parties involved in executing the post- disaster housing reconstruction project.
8. CBO's/General public	Assisted the affected community by way of supplying provisions such as food, clothing and other necessities at the immediate relief stage.

Source: Researcher's compilation

It is clear that post-disaster housing reconstruction is a co-ordinated effort for a unitary outcome. Hence, it is revealed that conflict of interests become explicit once the effort is operationalised, leading to expectation gaps. Therefore, classification of stakeholders is crucial in housing reconstruction projects.

## B. Classification of stakeholders in the context of post-disaster housing reconstruction

Stakeholder engagement in different capacities forms the basis of finding a classification of stakeholders in a post-disaster housing reconstruction scenario. The literature suggests that the affected community is the Primary Stakeholder (PS) and all the other parties including the government, private and professional organisations, NGO's, INGO's and CBO's are the Secondary Stakeholders (SS). During the pilot interviews, respondents were asked to classify the stakeholders. Table 3 presents the results of this classification.

**Table 3** Different classifications of stakeholders

Identified Stakeholders	Classification			
	Primary	External	Non-voluntary	Social
Affected community/ End user/ Beneficiary	Primary	External	Non-voluntary	Social
Consultant	Secondary	Internal	Non-voluntary	Non social
Contractor/sub contractors	Secondary	Internal	Non-voluntary	Non social
NGOs	Secondary	Internal	Non-voluntary/ Voluntary	Social
Donor/ Funding body	Secondary	Internal	Voluntary	Social/ Non social
Government	Secondary	Internal	Non-voluntary	Social/ Non social
CBOs representing Local people/general public	Secondary	Internal/ External	Non-voluntary	Social

Source: Researcher's compilation

Respondents held differing viewpoints in relation to categorisation of CBOs under the classification of internal-external, NGOs under the classification of voluntary- non-voluntary, donor agency and the government under social and non-social classifications. These variations were mainly due to the type of engagement and perceptual differences. It is also vital to identify the expectation gaps of the stakeholders in post-disaster housing reconstruction as it has formed the bases for the above compilation. However, in the aftermath of early reconstruction efforts, dissatisfaction was expressed by some stakeholders. Areas of criticism included time, cost, quality, coordination with infrastructure and linkage to livelihoods.

## C. Stakeholder expectations and gaps identified in post-disaster housing reconstructions in Sri Lanka

Stakeholder expectations and gaps are identified under three main titles: problems relating to quality and design of the houses, problems relating to relocation and institutional problems. The following sections demonstrate these problems in detail.

### (i) Problems related to time, cost, quality and design of the houses

Respondents suggested that timely handover of a housing project, keeping within the budget while retaining quality standards, and appropriateness to local cultural and social design, are crucial in post-disaster housing reconstruction. They also agreed that a gap in one or more criteria may lead to or aggravate the problem in some other area. Delays in plans for land acquisition and permanent housing in Sri Lanka had delayed timely completion of houses. Six interviewees expressed the view that a lack of strategies and resources to manage unanticipated situations such as escalation of construction costs, unanticipated funding gaps, material and labour shortages, amounted to delays in timely hand over of the housing

projects. One interviewee representing a private sector multinational organisation in Sri Lanka had experience of a project where the donor was the parent company and the project management team formed the employees in the organisation. The community affected formed the labour force, which helped to overcome resource constraints and enabled a smoother functioning of the process, even in the difficult situation.

All seven interviewees stated that nothing had been reported so far regarding the poor structures of the houses, and that they had used good quality materials. Also, the interviewees represented INGO's were locals who are based in Sri Lanka. However, interviewees were of the opinion that due to the shortages of building materials, labour, and other resources, the quality of housing had been affected. Moreover, an internal competition among the INGOs to complete the projects before the other, had overlooked the quality and design of housing. Moreover, the requirement of the beneficiary and the social norm had been poorly considered by the INGOs when designing the houses. Thus, they agreed that certain INGOs used materials which were of poor quality and did not suit the weather conditions of the country. Therefore, it is clear that the question of the sustainability of post-disaster housing is yet to be resolved.

## **(ii) Problems of relocation**

### **(a) Livelihood problems**

An interviewee representing the private sector and a donor agency revealed that their company looked into livelihood matters in collaboration with a local NGO. The victims were given masonry training where they could render the labour to generate an income, wherever they are within the homeland. Another interviewee disclosed that housing projects which were 'owner driven', where the victims occupying the house contributed to designing and building the house, reported fewer livelihood problems. However, relocation leading to livelihood problems should be addressed by relevant authorities such as the government. The interviewee representing the contracting organisation revealed that their main concern is profit but the organisation contributed more in discharging social responsibility when it came to Tsunami housing. Respondents also expressed their views on employment generation for the women who had been victimised twice by relocation and loss of their husbands. They expressed the view that it was a challenging task because many of the victimised women were housewives and the men had been the bread-winners. If they are to be given jobs then they should be trained or given resources for self employment which is an additional burden to the organisation. They also revealed that the post-disaster housing reconstruction budget did not provide any room to expand activities to cover such elements. One interviewee representing the consultancy organisation highlighted that, some of the youths who were previously engaged in temporary jobs, have now been depending on the financial assistance given by the government and a tendency for addiction to drugs and alcohol has been observed. Thus unemployment has burdened the economy, deteriorating the quality of life of the people.

### **(b) Mismatch between cultural and social patterns/norms**

An interviewee, representing the private sector, revealed that his organisation carried out a survey with the assistance of the government and a local NGO to identify the victims who came from the same neighbourhood, before deciding the place of relocation in the districts of Trincomalee and Batticaloa. Nevertheless, the interviewee representing the INGO expressed the view that they had carried out the construction where they were instructed to without considering the incompatibilities that might crop up in the future. Further, design of the houses has not catered for privacy issues, space requirements and cultural and social patterns.

Two interviewees agreed that their organisation could have given thought to local customs including the needs of extended families, the location of the kitchen, the building of verandas, privacy and weatherproofing when the houses are put up in the areas away from the original places. Nevertheless, they could not achieve these objectives due to the strict time limits, given design and location. Also, they expressed the opinion that the organisation did not

anticipate the behavioural outcome of the fisheries community, who used to live in individual houses, when being transferred to condominium properties. Social tension between the victims and the neighbourhood, too, has produced room for disharmony.

### (c) Infrastructure and other facilities

Six interviewees highlighted that infrastructure and other facilities were not taken into account when the post-disaster housing project was undertaken. They were also of the view that the government should take the sole responsibility of providing the infrastructure and facilities. One interviewee agreed that the project was executed within a strict time limit and that scarcity of resources such as materials, labour, lands and finances had hindered the consideration of the related infrastructure and other facilities attached to the house. Moreover, insufficient infrastructure and other inadequate services such as water and sanitation, electricity, transportation systems, postal services, education facilities, medical facilities etc. have aggravated victims' traumatic situation. According to Lyons (2009), the following were the own words of a victimised resident, who got relocated:

*"No transport. . . no postal service . . . and there is no system in place for such activities. . . the absence of street lamps—in some seasons elephants come . . .and with the absence of lighting it is very dangerous to travel on the roads in the night. . . The village has a huge garbage problem."*

The above statement affirms the importance of a proper infrastructure system and other services which are interlinked to the efforts of post-disaster housing reconstruction. It is observed that due to poor access to infrastructure and other services, several other problems including degradation of quality of life have cropped up in Sri Lanka. Thus, post-disaster housing reconstruction has many other aspects to be looked into apart from 'providing a house'.

### **(iii) Institutional problems**

Two interviewees expressed the opinion that the objectives of their project were fully achieved. For example, one respondent described an occasion where an external audit was carried out to ensure that funds allocated for the housing project were not misappropriated. Also, they revealed that communication and co-ordination with other stakeholders was well in place during the process of reconstruction. They also revealed that the projects undertaken in the north and east of the country were delayed due to the ongoing civil war, where they had to work in consultation with the members representing the group who fought for a separate homeland in the north and the east.

Post-disaster housing reconstruction undertaken by the INGO has outsourced the building part to a private contractor where participation of the community is minimal. Also, communication with the government and consultation of the community and other authorities had been carried out at a remote level, which has barred essential local knowledge and suitable policies from coming into the process of reconstruction. This might lead to the creation of confusion among stakeholders, and development of a unitary policy framework is inhibited.

One respondent revealed that the structure of the organisation is such that spontaneous decisions were suppressed. The process of decision making had to go through a cumbersome process which took a prolonged time. For example, obtaining the title deed of the lands that were outside the buffer zone became a time consuming task due to the red tape and bureaucratic structures. All seven interviewees agreed that decision making in relation to the reconstruction process had lagged behind due to the political uncertainty, exacerbated by the civil conflict and the tendency of incoming governments to overturn the arrangements of their predecessors, which contributed to a poor investment climate.

Though key challenges in relation to post-tsunami housing resettlement were jointly identified by stakeholders, their expectations were poorly addressed.

## 5. Discussion

Stakeholder groups and their engagements identified in the literature and interviews were similar. Thus, the literature enumerates stakeholders in construction project contexts in a concrete manner, and the interviewees held different viewpoints in their classifications. This confirms that no one person perceives the world in the same way, which is due to many factors. Environment, social standing, exposure, education, experience, culture, attitudes, values are some of the obvious reasons. Thus the world is subject to change. Moreover, the emergence of two types of housing reconstruction projects, 'owner driven' and 'donor driven', had an impact on these classifications. Therefore, the factors mentioned above too are subject to change. Nevertheless, the interview findings highlighted that the parties engaged in post-disaster housing reconstruction have not diverted attention on providing related infrastructure.

Issues in relation to time, cost and problems of relocation which have been unearthed in the literature and interviews reflected similar results, except for one occasion. There was evidence that, if the supply chain partners are within the organisation, it will minimise the adverse effects of macroeconomic factors such as inflation. One respondent revealed that the organisation he represented carried out a post-project audit to ensure the transparency and accountability of transactions. Communication and coordination with stakeholders and other groups, community participation, timely decision making etc. were noted as important ingredients of post-disaster housing reconstruction by all the interviewees.

The literature provided a general overview of expectation gaps specific to post-disaster housing reconstruction in Sri Lanka, whereas interview results revealed the problems specific to personal experience and to that of the organisations. Table below is a summary of key success / failure factors identified from literature and interview findings.

**Table 4** Key Success/ failure factors of post-disaster housing

<b>Success/ failure factors</b>
1. Time a. Timely supply of material and labour (expert opinion and construction craftsmen) b. Timely handover
2. Cost a. Meeting the budget b. Low cost building materials, labour and other resources
3. Quality a. Quality standards of materials, labour
4. Features of the house a. Location b. Design of the house suiting to cultural and social patterns c. Space d. Location e. Changeability f. Aesthetics
5. Availability of resources- Physical and human (Professional/ expert advice, information)
6. Unitary policies and frameworks



7.	Supervision
8.	Coordination, consultation and collaboration
9.	Supportive laws and regulations
10.	Clearly defined goals and commitments
11.	Effective information management
12.	Effective communication
13.	Sufficient mobilisation and disbursement of resources
14.	Value addition to the product by introducing new technologies and materials
16.	Community participation
17.	Employment generation
18.	Availability of infrastructure and other facilities
19.	Post completion audit

Source: Researcher's compilation

The above discussion is based on the findings of the literature and seven interviews followed by an open-ended questionnaire. The sources of gathering data are limited only to two, a literature review and synthesis and interviews. The outcomes are also limited to only two sources. Therefore, it is necessary to carry out case study interviews with experts in post-disaster housing reconstruction projects, representing the government, the private sector, NGO's, INGO's, CBO's and humanitarian organisations in Sri Lanka, in order to get a panoramic view of stakeholder expectations.

## 6. Conclusion

Post-disaster housing reconstruction projects are mostly community based and non-profit oriented in contrast to normal construction projects. The stakeholders of such a project could be defined as any individual or a group who are affected and / or who could affect the achievement of the post-disaster housing reconstruction project objectives. The identification of stakeholders suggested by the literature and the findings of the pilot study remained the same. Interviewees' categorisation of stakeholders under different classifications varied apropos to internal-external, voluntary-non voluntary and social-non social classifications. Differing classifications emerged due to the environment, social standing, exposure, education, experience, culture, attitudes, values and different types of housing reconstruction (owner and donor driven) projects that some of the interviewees got exposed to. Stakeholder expectations and gaps identified in the literature and in real life scenarios represented similarities. Remedial action for livelihood problems was taken occasionally. Infrastructure and other services were least addressed. Community participation, meeting budgetary and time requirements, delivery of a high quality, value added product are perceived as criteria for improving stakeholder expectations. Hence, participative decision making, proper planning, transparency, and accountability integrated with good communication and coordination, could all reduce the gaps identified apropos to design and quality, relocation and institutions in post-disaster housing reconstruction in Sri Lanka. This paper sets out a few preliminary findings of the literature survey and pilot interviews where the opinions of individuals are highlighted. Detailed case study interviews have to be carried out to gain a holistic picture of the scenario. The researcher expects to carry out case study interviews with experts representing the government, private sector, NGO's, INGO's, CBO's and humanitarian organisations in the future.

## References

- Andriof, J. and Waddock, S. (2002) *Unfolding stakeholder engagement*, Greenleaf Publishing, 2002.
- Ariyabandu, M.M. and Hulangamuwa, P. (2002) *Corporate Social Responsibility and Natural Disaster Reduction in Sri Lanka*, ITDG South Asia, Sri Lanka.
- Bakker, F.G.A. de and Hond, F. den (2008) Introducing the Politics of Stakeholder Influence A Review Essay , *Business and Society*, Vol.47(1), pp. 8-20.
- Bhattacharjee, A., Rawal, V., Fautin, C., Moore, J.-L., Kalonge, S., & Walden, V. M. et al. (2005) *Multi-Agency Evaluation of Tsunami Response: India and Sri Lanka*. CARE International, Oxfam GB and World Vision International.
- Billa, L., Shattri, M., Mahmud, A.R. and Ghazali, A.H. (2006) Comprehensive planning and the role of SDSS in flood disaster management in Malaysia, *Disaster Prevention and Management*, Vol. 15 (2), pp. 233-240.
- Boano, C. (2009) Housing anxiety and multiple geographies in post-tsunami Sri Lanka, *Disasters*, Vol.33 (4), pp. 762-785.
- Broadbent, S.G and Broadbent, D.M. (2006) Rethinking Future UK Support to Post-disaster Reconstructions : Meeting stakeholder interests, Proceedings of a Conference Held at the Scuola di Sanità Militare, Florence, Italy, 17-19 May 2006.
- Bunn, M.D., Savage, G.T., Holloway, B.B. (2002) Stakeholder analysis for multi-sector innovations, *Journal of Business and Industrial Marketing*, Vol.17 No.2/3, pp. 181-203.
- Carroll, A.B. (1999) "Corporate social responsibility", *Business and Society*, Vol. 38 No.3, pp.268-95.
- Christian Aid (2003) "Part one: A dangerous world", in Palakudiyil, T. and Todd, M., *Facing up to the Storm: How Local Communities Can Cope with Disaster: Lessons from Orissa and Gujerat*, 1st ed., Christian Aid Publications, London (Available at: [www.christian-aid.org.uk/indepth/0307stor](http://www.christian-aid.org.uk/indepth/0307stor) [assessed 20/02/2008]).
- CIOB (2007) Corporate Social Responsibility and Construction Summary (Available online at [http://www.google.co.uk/search?sourceid=navclient&hl=en-GB&ie=UTF-8&rlz=1T4GGLL\\_en-GBGB313GB313&q=CIOB+stakeholdes](http://www.google.co.uk/search?sourceid=navclient&hl=en-GB&ie=UTF-8&rlz=1T4GGLL_en-GBGB313GB313&q=CIOB+stakeholdes) [Accessed 23/04/2007])
- CRED (2007) "The OFDA/CRED international disasters database", (Available online <http://www.cred.be/emdat/disdat.htm>. [Accessed on 20/10/2007]).
- CRED (2008) "The OFDA/CRED international disasters database", (Available online <http://www.cred.be/emdat/disdat.htm>. [Accessed on 02/08/2008]).
- Davidson, C.H, Johnson, C, Lizarralde, G., Dikmen, N., and Sliwinski, A. (2007) Truths and myths about community participation in post-disaster housing projects, *Habitat International*, Vol.31, pp. 100-115.
- Department of Census and Statistics(2005) "Impact of 2004 Tsunami on Sri Lanka", (Available at [www.statistics.gov.lk/Tsunami/index.htm](http://www.statistics.gov.lk/Tsunami/index.htm). [Accessed on 16/08/2007]).
- Fazlulhaq, N. (2008) Deadline extended for tsunami housing projects, *The Sunday Times* 12 December 2008.
- Global Forum for Disaster Reduction (GFDR) and Graduate School of Global Environmental Studies (2005), *The Corporate Sector Role in Disaster and Environmental Management: beyond Corporate Social Responsibility*, White Paper on Corporate-Community Interface (CCI), Kyoto University.
- GoSL (2005) "*Sri Lanka: Post Tsunami Recovery and Reconstruction*", Joint Report of the Government of Sri Lanka and Development Partners, December.
- GoSL (2006) *Post Tsunami Recovery and Reconstruction*, Joint Report of the Government of Sri Lanka and Development Partners. December 2006.
- Hayles, C. (2008) An exploration of current planning, design and building issues in post-disaster housing reconstruction, *Proceedings of the CIB International Conference on*

- Building Education and Research, 11th-15<sup>th</sup> February, 2008, Kandalama, Sri Lanka, pp.180-181.*
- Haigh, R. , Amaratunga, D. and K. Kerimanginaye (2006) An Exploration of the Construction Industry's Role in Disaster Preparedness, Response and Recovery, *The construction and building research conference of the Royal Institution of Chartered Surveyors, 7-8 September 2006, London, University College London.*
- Hopkins, M. (2005) *Business, CSR and the Tsunami: Time to re-think corporations and development?* (available online at <http://www.mhcinternational.com/articles/csr%20development%20and%20the%20tsunami.htm> [Accessed on 21 /02/2008]).
- Hoyois, P., Below, R., Scheuren, J.M. and Guha-Sapir, D. (2007) *Annual Disaster Statistical Review: Numbers and Trends 2006*, CRED, Brussels.
- ICRC (2006) Operational Update, ICRC Bulletin No. 04 / 2006- Sri Lanka (Available online at <http://www.icrc.org/web/eng/siteeng0.nsf/htmlall/srilanka-news-100806?opendocument>[Accessed on 20/10/2007]).
- Jigyasu, R. (2002), "From Marathwada to Gujarat – emerging challenges in post earthquake rehabilitation for sustainable eco-development in South Asia", *Proceedings of the First International Conference on Post-disaster Reconstruction: Improving Post-Disaster Reconstruction in Developing Countries, 23-25 May, Universite de Montreal, Canada.*
- Jigyasu, R. (2004) Sustainable post disaster reconstruction through integrated risk management – the case of rural communities in South Asia, *Proceedings of the 2004 International Conference and Student Competition on post-disaster reconstruction "Planning for reconstruction", 22-23 April, Coventry, UK.*
- Keraminiyage, K., Amaratunga, D. and Haigh, R. (2006), Role of construction in managing disasters in developing economies, The Research Institute for the Built and Human Environment ,The University of Salford ,Salford, UK.
- Kurita, T., Nakamura, A.,Kodama, M. and Colombage, S.R.N. (2006) Tsunami public awareness and the disaster management system of Sri Lanka, *Disaster Prevention and Management*, Vol. 15 (1), pp. 92-110.
- Lyons, M. (2009) Building Back Better: The Large-Scale Impact of Small-Scale Approaches to Reconstruction, *World Development* Vol. 37(2), pp. 385–398.
- Mitchell, R.K., Agle, B. R., Wood, D.J. (1997) Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What Really Counts, *The Academy of Management Review*, Vol. 22, No. 4, pp. 853-886.
- Mulligan, M. and Shaw, J.(2007) What the world can learn from Sri Lanka's post-tsunami experiences, *International Journal of Asia Pacific Studies* Vol. 3(2), pp.65-91.
- Nissanka, N.M.N.W.K., Karunasena, G. and Rameezdeen, R. (2008) Study of factors affecting Post-Disaster Housing Reconstruction, *Proceedings of the CIB International Conference on Building Education and Research, 11th-15<sup>th</sup> February, 2008, Kandalama, Sri Lanka, pp 186-187.*
- Newcombe R. (2003) From client to project stakeholders: a stakeholder mapping approach, *Construction Management and Economics*, Vol. 21, pp. 841–848.
- Olander, S. and Landin, A. (2005) Evaluation of stakeholder influence in the implementation of construction projects, *International Journal of Project Management* , Vol. 23 (2005), pp. 321–32.
- Oloruntoba, R. (2005) A wave of destruction and the waves of relief: issues, challenges and strategies, *Disaster Prevention and Management*, Vol. 14(4), pp. 506-521.
- Oxfam International. 2005 "Oxfam International Tsunami Fund 4th Quarter Report: October to December 2005," (Available at [www.oxfam.org/en/files/oitf\\_report\\_2005q4/download](http://www.oxfam.org/en/files/oitf_report_2005q4/download) (Accessed on 25/11/ 2006).
- Pardasani, M. (2006) Tsunami reconstruction and redevelopment in the Maldives A case study of community participation and social action, *Disaster Prevention and Management*, Vol. 15(1),pp.79-91.

- Peek, L.A. and Sutton, J.N. (2003) An Exploratory Comparison of Disasters, Riots and Terrorist Acts, *Disaster*, Vol. 27, No.4, pp 319-335.
- Perry, M. (2007) Natural disaster management planning: A study of logistics managers responding to the tsunami, *International Journal of Physical Distribution & Logistics Management*, Vol. 37(5), pp. 409-433.
- Podnar, K. and Jancic, Z. (2006), Towards a Categorization of Stakeholder Groups: An SUM Consult GmbH, Wiesbaden, Germany (2005) Efforts in South-East Asia Assessment Study for the Preparation of the Second Phase of Funding, Report submitted to the Austrian Development Cooperation, Vienna, Austria.
- ProVention Consortium and ALNAP (2005) South Asia Earthquake 2005 Learning from previous recovery operations, *ProVention publications*, 2005.
- Ratnayake, R.M.G.D and Rameezdeen, R. (2008) A comparative study of donor driven vs. owner driven post disaster housing reconstruction, *Proceedings of the CIB International Conference on Building Education and Research, 11th-15<sup>th</sup> February, 2008*, Kandalama, Sri Lanka, pp 192-193.
- Ratnayake, R.M.G.D and Rameezdeen, R. (2008) A comparative study of donor driven vs. owner driven post disaster housing reconstruction, *Proceedings of the CIB International Conference on Building Education and Research, 11th-15<sup>th</sup> February, 2008*, Kandalama, Sri Lanka, pp 192-193.
- ReliefWeb (2008) Sri Lanka: Tsunami victims threatened with homelessness, (Available online at <http://www.reliefweb.int/rw/rwb.nsf/db900SID/FBUO-7K7GGF?OpenDocument> [Accessed on 23/11/2008])
- RICS, 2006, *Mind the gap! Post-disaster reconstruction and the transition from humanitarian relief*, (UK: Royal Institution of Chartered Surveyors), Westminster.
- Rex, S. (2006), "Transforming organisations: from development organisations to disaster response programme: a case study in capacity building". *Proceedings of the First International Conference on Post-disaster Reconstruction: Meeting Stakeholder Interests*, 17-19 May, Florence, Italy.
- SUM Consult GmbH, Wiesbaden, Germany (2005) Efforts in South-East Asia Assessment Study for the Preparation of the Second Phase of Funding, Report submitted to the Austrian Development Cooperation, Vienna, Austria.
- Sundnes, K.O and Birnbaum, M.L. (2002) *Health Disaster Management- Guidelines for Evaluation and Research in the Utstein Style*, Healthcare Fields Intensive Care and Emergency Medicine, University of Wisconsin.
- TAFREN (2005) *Housing and township development: Assistance policy and implementation guidelines*. Colombo: TAFREN.
- TEC (Tsunami Evaluation Coalition) (2006a) Impact of the Tsunami on local and national capacities [Scheper, E., Parakrama, A., & Patel, S., with contributions from Vaux, T.].
- Telford, J., Cosgrave, J. and Houghton, R. (2006) *Joint Evaluation of the international response to the Indian Ocean tsunami: Synthesis Report*. London: Tsunami Evaluation Coalition. London.
- Twigg, J. (2002) *Corporate Social Responsibility and Disaster Reduction- Conclusions and Recommendations*, Benfield Greig Hazard Research Centre, University College London.
- UNDP (2005) *A Global Review: UNDP Support to Institutional and Legislative Systems for Disaster Risk Management*, Executive Summary of World Conference on Disaster Reduction, Kobe.
- UN/ISDR (2002) *Living with risk: a global review of disaster reduction initiatives*, Preliminary version prepared as an interagency effort coordinated by the ISDR Secretariat, Geneva.
- UN/ISDR (2005) *Hyogo Framework for Action 2005-2015: Building the Resilience of Nations and Communities to Disasters*, World Conference on Disaster Reduction 18-22 January 2005, Kobe, Hyogo, Japan.
- UN/ISDR (2008) *Disaster figures for 2007*, United Nations, secretariat of the International Strategy for Disaster Reduction, Geneva.

- UN/OCHA (2005) Indian Ocean Earthquake- Tsunami 2005 Mid Term Review (Available online at <http://ochaonline.un.org/humanitarianappeal/webpage.asp?Page=1284#Regional> [Accessed on 26/08/2006]).
- Walker, D.H.T. (2000) Client/Customer or Stakeholder Focus? ISO 14000 EMS as a construction industry case study, *The TQM Magazine*, volume 12(1), 18-25
- Walker, D.H.T., Bourne, L.M. and Shelley, A. (2008) Influence, stakeholder mapping and visualization, *Construction Management and Economics*, Vol. 26, pp. 645-658.
- Weerakoon, D., Jayasuriya, S., Arunatilake, N. and Steele, P. (2007) *Economic Challenges of Post-Tsunami Reconstruction in Sri Lanka*, ADB Institute Discussion Paper No. 75.
- Winch, G.M. (2002) *Managing Construction Projects: An Information Processing Approach*, Blackwell Science Ltd., Oxford.
- World Bank (2006) *Hazards of Nature, Risks to Development: An IEG Evaluation of World Bank Assistance for Natural Disasters*, World Bank, Washington D.C.
- World Health Organisation (2006) , Collaboration within the United Nations system and with other intergovernmental organizations, including United Nations reform process, *Report by the Secretariat*, 59<sup>th</sup> Assembly, Geneva, Switzerland.

# Managing Disaster in Indonesia

**Taufika Ophiyandri, Dilanthi Amaratunga and Chaminda Pathirage**, *Research Institute for the Built and Human Environment (BuHu), University of Salford*

## Abstract

Indonesia is prone to disaster, be it natural or man-made. Some causes of these disasters are related to geography, geology, climate or other factors related to social, cultural or political diversity. Its location at a juncture of four tectonic plates, make it very vulnerable to earthquakes. The giant earthquake which triggered the tsunami on 26 December 2004 was the most devastating disaster in Indonesia. It took 130,000 people, displaced 500,000 people, and destroyed 120,000 houses. This disaster highlighted the importance of disaster management. In this context, this paper tries to detail the problems faced by Indonesia in managing and implementing disaster management. The paper is based on a detailed literature review.

## 1. Introduction

The Republic of Indonesia is a country in Southeast Asia. It is the world's largest archipelagic state that consist of 17,508 islands, in which 6000 are inhabited (NIA, 2004). The estimated population in 2008 was 237 million. Figure 1 shows the political map of Indonesia.



**Figure 1.** Indonesia political map (source USGS, 2008a)

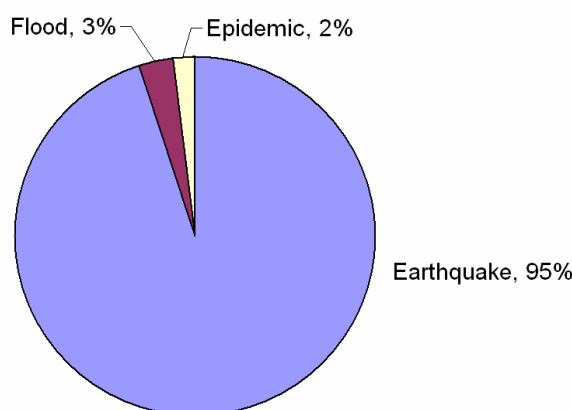
Lying along the equator, Indonesia has a tropical climate, with two distinct monsoonal wet and dry seasons. Average annual rainfall in the lowlands varies from 1,780–3,175 millimeters, and up to 6,100 millimeters in mountainous regions. Mountainous areas—particularly in the west coast of Sumatra, West Java, Kalimantan, Sulawesi, and Papua—receive the highest rainfall. Humidity is generally high, averaging about 80%. Average temperatures are classified as follows: coastal plains: 28°C; inland and mountain areas: 26°C; higher mountain areas: 23°C (NIA, 2004).

Indonesia is a beautiful country, but also prone to disasters, natural or man-made. Some causes of disaster are related to geography, geology, climate or other factors related to social, cultural or political diversity. The most devastating disaster was the boxing day tsunami on 26 December 2004. The Agency for the Rehabilitation and Reconstruction Aceh and Nias (BRR) (2006) reports that in Aceh 130,000 people were confirmed dead, 37,000 missing, and 500,000 were displaced. This disaster has become a wake-up call for Indonesians on how to face disasters. Since then, there has been a new paradigm on facing disaster in Indonesia, from emergency response to risk management. However, there are a lot of problems with implementation. Hence, this paper tries to detail the problems faced by Indonesia in managing and implementing disaster management which based on a detailed literature review.

## 2. Type of disasters

According to the United Nations International Strategy for Disaster Reduction (UN-ISDR) (2008), disaster is a serious disruption of the functioning of a community or a society causing widespread human, material, economic or environmental losses which exceed the ability of the affected community or society to cope using its own resources. It is a function of the risk process which results from the combination of hazards, conditions of vulnerability and insufficient capacity or measures to reduce the potential negative consequences of risk (UN-ISDR, 2008). Shaluf (2008) classifies disasters into three types: natural disasters are those disasters that result from natural forces, man-made disasters are those disasters that result from human decisions, and hybrid disasters are those disasters that resulted from both natural and man-made causes.

Disasters in Indonesia have increased by the year, especially in the last decade. The National Development Planning Agency (BAPPENAS) and National Coordinating Board for Disaster Management (BAKORNAS PB) counted, within the period of 2003-2005 alone, 1,429 disaster incidences with hydro-meteorological disasters comprising the bulk at 53.3 percent (BAPPENAS and BAKORNAS PB, 2006). Of this figure, floods occur most often (34.1percent), followed by landslides at 16 percent. Although geological disasters (earthquakes, tsunamis and volcanic eruptions) comprise only 6.4 percent of the total, they caused tremendous loss and fatalities. Most fatalities were accounted for by the combined earthquake and tsunami disaster in the Province of Nanggroe Aceh Darussalam and North Sumatra on 26 December 2004 and the massive earthquake in Nias, North Sumatra on 28 March 2005. In Indonesia, earthquakes took the highest percentage (95%) of reported people killed by disaster in Indonesia from 1980-2008 (Figure 2). During this time, the top five deadliest disasters were also earthquakes (Table 1) (PreventionWeb, 2008).



**Figure 2.** Percentage of reported people killed by disaster type in Indonesia from year 1980-2008 (source: PreventionWeb, 2008)

**Table 1.** Top 10 Natural Disaster reported – killed in Indonesia from 1980-2008 (source: PreventionWeb, 2008)

Disaster	Date	Killed (no. of people)
Earthquake*	2004	165,708
Earthquake*	2006	5,778
Earthquake*	1992	2,500
Earthquake*	2005	915
Earthquake*	2006	802
Epidemic	1998	777
Drought	1997	672
Epidemic	1998	672

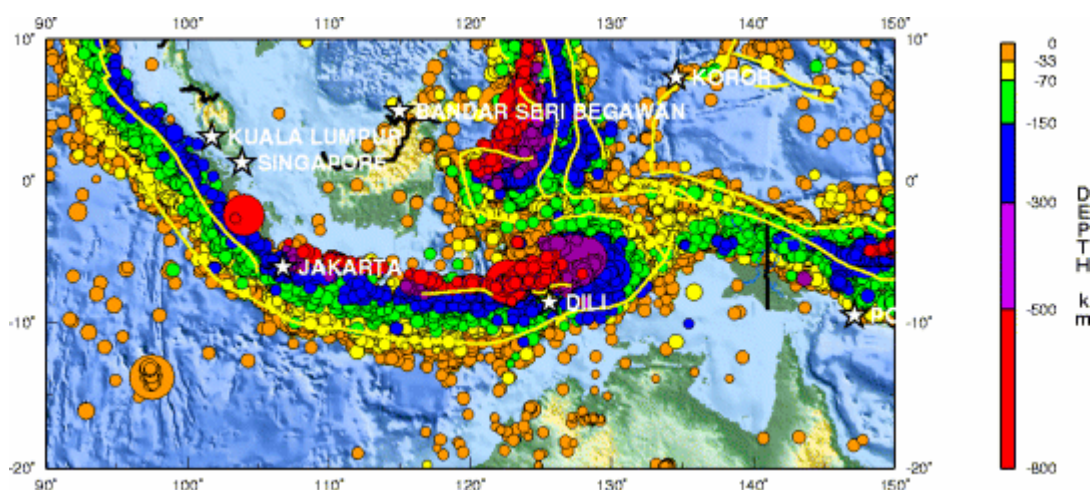
Epidemic	2004	658
Flood	1981	500

Below are the details of information on types of disaster in Indonesia which are mainly extracted from BAPPENAS and BAKORNAS PB (2006).

### 2.1. Earthquakes and tsunamis

Earthquakes occur relatively frequently in Indonesia because of tectonic plate movements and volcanic eruptions. Tectonic plate movements often occur along Sumatra's west coast where the Asian Plate adjoins the Indian Ocean Plate; Java's south coast and the islands of Nusa Tenggara where the Australian Plate meets the Asian Plate; and Sulawesi and Maluku where the Asian Plate meets the Pacific Ocean Plate. They produce an earthquake belt in Indonesia with thousands of epicentres and hundreds of hazardous volcanoes (BAPPENAS and BAKORNAS PB, 2006).

Figure 3 shows the seismicity map of Indonesia from 1990-2006 (USGS, 2008b). The intensity of big earthquakes has risen significantly since 2000. According to USGS (2008c), from the year 1900 until 2000 only 6 devastating earthquakes happened, but after 2000, the number of occurrences has increased by 38.



**Figure 3.** Indonesia seismicity map, 1990-2006 (source USGS, 2008b)

Tectonic plate movement-induced earthquakes that happen under the sea often generate tidal waves. Highly predisposed to tectonic plate movement, Indonesia often experiences tsunamis. Tsunamis in Indonesia are mostly generated by tectonic earthquakes occurring along subduction zones and other seismic active areas. Between 1600 and 2000 there were 105 tsunamis of which 90 percent were generated by tectonic earthquakes, 9 percent by volcanic eruptions and 1 percent by landslides. Indonesia's coastal areas are prone to tsunamis. These areas include the west coast of Sumatra, the south coast of Java, the north and south coast of Nusa Tenggara, the Maluku islands, the north coast of Irian Jaya and the greater part of the coast of Sulawesi. The Maluku Sea is the most tsunami-prone area. Between 1600 and 2000 it experienced 32 tsunamis of which 28 were generated by earthquakes and four by undersea volcanic eruptions (BAPPENAS and BAKORNAS PB, 2006).

Earthquakes and tsunamis normally inflict devastating loss of lives and assets, and create a need for lengthy rehabilitation and reconstruction. The tsunami which triggered the most immense and widespread destruction in modern history took place in the Indian Ocean following a 8.9 Richter scale earthquake nearby Simeulue, the Province of Nanggroe Aceh Darussalam, on 26 December 2004. It devastated Banda Aceh City, Aceh's west coast and Nias Island. The disaster also wreaked havoc in countries along the Indian Ocean including Thailand, Malaysia, Andaman and Nicobar, Sri Lanka and even the coast of East Africa. The



total economic loss in Indonesia is estimated at 41 trillion rupiah, exclusive of indirect losses caused by interrupted productive and economic activities (BAPPENAS and BAKORNAS PB, 2006). On 28 March 2005, another tragedy struck when an earthquake hit off the coast of Sumatra, near the island of Nias. According to BRR (2006) nearly 900 people died as a result, the majority of buildings and infrastructure suffered damage, leaving some 40,000 displaced.

The Aceh and Nias disasters have devastated communities and created massive damage to infrastructure, schools, hospitals, the environment and the economy. BRR (2006) reports that housing requirements were more than 120,000 new houses and rehabilitation of almost 85,000, both medium and heavily damaged homes. This sector suffered almost USD 1.4 billion in damages, which accounted for over 30% of the overall damage and losses.

While Indonesia was still recovering from the devastating impact of the tsunami and earthquake in Aceh and Nias in 2004, in 2006 it was confronted with another strong earthquake and tsunami. On 27 May 2006, an earthquake struck the very heartland of Indonesia, near the historic city of Yogyakarta. With its epicenter in the Indian Ocean at about 33 kilometers south of Bantul district, it measured 6.3 on the Richter Scale and lasted for 52 seconds. Because the earthquake was relatively shallow at 33 kilometers under ground, shaking on the surface was more intense than deeper earthquakes of the same magnitude, resulting in major devastation, in particular in the districts of Bantul in Yogyakarta Province and Klaten in Central Java Province. The Java Reconstruction Fund (JRF) (2007) reports that it claimed more than 5,700 lives and destroyed over 280,000 homes. The total damage and losses were estimated to be around Rp 29.1 trillion (US\$3.1 billion).

Just two months later, on 17 July 2006, a second major submarine earthquake of magnitude 7.7 on the Richter scale struck off the southern coast of Java, creating a three meter-high tsunami that displaced over 28,000 people and taking more than 650 lives. The worst impact of the tsunami was felt in the district of Ciamis, West Java, although damage was suffered as far as Central Java to the east. Along the coast of Ciamis close to 6,000 families were displaced. In total over 1,900 houses were destroyed and 514 heavily damaged. Some 1,200 trade facilities and other public facilities were destroyed. Estimates of damage and losses in the tsunami affected areas were Rp 1.004 trillion (US\$110.3 million) (JRF, 2007).

## *2.2. Volcanic eruptions*

Presently there are 129 active volcanoes and 500 inactive volcanoes in Indonesia. Of the total active volcanoes, which account for approximately 13 percent of the world's total number, 70 are eruptive and 15 are critical (BAPPENAS and BAKORNAS PB, 2006).

The volcanoes form a belt covering, in one straight stretch, Sumatra, Java, Bali and Nusa Tenggara, before making a northwards turn to the Banda Sea and the northern part of Sulawesi. The belt's length is about seven thousand kilometres and it contains volcanoes of mixed characteristics. Presently more than 10 percent of Indonesia's population inhabit regions prone to volcanic eruptions. Over the last 100 years more than 175,000 have become victims of volcanic eruptions (BAPPENAS and BAKORNAS PB, 2006).

Indonesia experiences dry and rainy weather as it lies in the tropical climate zone. Therefore, aside from pyroclastic avalanches, volcanic eruptions also pose indirect hazards such as lahar flows or hazardous volcanic material movements (BAPPENAS and BAKORNAS PB, 2006).

## *2.3. Floods*

Floods reoccur annually in Indonesia, most notably during the rainy season. From a morphologic point of view, floods happen on account of the country's highly varied landscape and its many rivers. Floods generally occur in Indonesia's western region, which experiences a heavier rainfall than the eastern region. Indonesia's growing population and need for space to accommodate life support activities has indirectly contributed to floods. Logging has been resorted to address the demand for space, to increase sedimentation into rivers which in turn produce uncontrolled runoff and high groundwater saturation. It may cause runoff to reach such an extent as to cause a flash flood, as happened in 2003 in the sub-districts of Bahorok and Langkat (North Sumatra) and in the sub-district of Ayah (Central Java), all of which claimed considerable numbers of lives and assets (BAPPENAS and BAKORNAS PB, 2006).

In 2006 sizeable disasters took place, including landslides and flash floods in East Java, Jember, where 92 were killed and 8,861 displaced and Trenggalek, where 18 were killed. A flash flood accompanied by a landslide occurred in Manado (North Sulawesi), claimed 27 lives and displaced 30,000. A similar disaster happened in South Sulawesi in June 2006. More than 200 were killed and tens remained missing (BAPPENAS and BAKORNAS PB, 2006).

#### *2.4. Landslides*

Landslides in Indonesia frequently occur in areas with steep slopes. Landslides are normally triggered by heavy rainfall. Data suggest that regions highly prone to landslides are Sumatra's Bukit Barisan mountains and mountain ranges in Java, Sulawesi and Nusa Tenggara. Fatal landslides also occur at drilling sites and mining shafts. Landslides also occur annually, in particular in regions with unstable land such as West Java and Central Java (BAPPENAS and BAKORNAS PB, 2006).

Most lands in tropical regions are prone to landslides because of high degrees of rock weathering, causing soil composition to be dominated by loose layers of material. Land stability is greatly influenced by damaged buffer zones in Indonesia. Extensive logging in buffer zones has only increased the potential for landslides. An example is West Java, which in 1990 registered 791,519 ha of forest areas (accounting for approximately 22 percent of its total area) only to have it reduced to 323,802 ha by 2002 (approximately only 9 percent of total area). It could only be expected that landslides occur frequently in this province (BAPPENAS and BAKORNAS PB, 2006).

#### *2.5. Drought*

If floods and landslides happen during the rainy season, droughts potentially occur during the dry season. Droughts pose an even more serious problem when impacting food crop-producing regions, as once happened in Bojonegoro which had 1,000 ha of its rice fields experience a harvest failure when the irrigation system dried up. A similar case happened at Java's north coast, where a drought struck 12,985 ha of food crop producing areas (BAPPENAS and BAKORNAS PB, 2006).

Nowadays, droughts also impact the country's power supply as they lower the output of power plants, in particular hydroelectric power plants catering to the energy needs of Java and Bali. Droughts normally occur during long dry seasons in some regions especially Indonesia's Eastern Regions such as NTB, NTT and several areas in Sulawesi, Kalimantan and Papua. Droughts also potentially promote the spread of tropical diseases such as malaria and dengue fever (BAPPENAS and BAKORNAS PB, 2006).

#### *2.6. Forest and land fires*

Forest fires happen at an increasingly alarming rate. The negative impacts of forest fires are considerable, including decreased biodiversity, reduced economic value of forests, declining land productivity, micro and global climate changes, and smoke hazes hazardous to health and disruptive to land, river, lake, sea and air transport. A considerable forest fire took place in Indonesia in 1997-1998 in Kalimantan and parts of Sumatra. Kalimantan and Sumatra are prone to land fire because the land there is composed of peat, which burns easily following uncontrolled land clearing (BAPPENAS and BAKORNAS PB, 2006).

Data from the Directorate General for Forest Protection and Conservation suggest that between 1998 and 2002 forest fires have occurred annually, each time impacting between 3,000 and 515,000 ha of forest (BAPPENAS and BAKORNAS PB, 2006).

Forest and land fires are mainly triggered by natural factors and by human activities such as land clearing. Low welfare and low education of communities living in and around forests are among the contributing factors. The more prominent factor contributing to forest fire is the logging activity conducted by forest concession holders who clear forests without heeding regulations and environmental considerations (BAPPENAS and BAKORNAS PB, 2006).

### *2.7. Epidemics, outbreaks of disease and extraordinary events*

Epidemics, outbreaks of disease and extraordinary events are hazards caused by the spread of communicable diseases in a certain region. On a large scale, these may cause fatalities and increase the number of infected persons. Diseases that are currently being watched rigorously in Indonesia include dengue fever, malaria, bird flu, anthrax, hunger oedema and HIV/ AIDS. Their spread is generally very hard to contain, hence what starts out as a local event may quickly turn into a national disaster claiming many lives. Poor environmental conditions, low quality foods and life styles and climatic changes are some of the contributing factors (BAPPENAS and BAKORNAS PB, 2006).

### *2.8. Technological disasters*

Failure of a technological system leading to a technological disaster can always be traced back to a system error, which in turn is caused by a system design that is incompatible with the environment it operates in - design that fails to strike a balance between the technical and the social systems. It happens frequently in Indonesia to fatal consequences. Examples include transportation accidents (involving ships, aircrafts, land transport such as trains), industrial accidents (gas pipe leaks, poisoning, environmental pollution) and household accidents (short-circuits, fires). Transgenic technological disasters also pose a potential threat considering that Indonesia is an open market to transgenic products (BAPPENAS and BAKORNAS PB, 2006).

Transportation accidents constitute the bulk of technological disasters in Indonesia, claiming on average 30,000 victims a year. In this category, Indonesia ranks third among ASEAN countries (BAPPENAS and BAKORNAS PB, 2006).

A technological disaster that is ongoing now is a blow-out on an oil well in Sidoarjo, East Java. The mud blow-out happened following a faulty prediction which neglected the characteristics of underground pressurized rock, causing water-saturated mud to flow continuously from the well to inundate business operations/rice fields and local farmlands and homes. The mudflow has also led to the closure of nearby factories and the Surabaya-Gempol toll road disrupting distribution of industrial goods. It is feared that massive losses and environmental damage will be sustained should the industrial accident not be immediately addressed (BAPPENAS and BAKORNAS PB, 2006).

### *2.9. Social unrest*

Indonesia's social culture, made up of diverse ethnicities, races, community groups, languages and religions, is a rich national asset and at the same time a highly vulnerable feature. Those with vested interests have often taken advantage of it to incite conflict. Indonesia's proneness to conflict is made worse by gaping economic disparities and low education. Also contributing is a declining sense of nationalism as demonstrated by some regions aspiring to break away from the Unitary State of the Republic of Indonesia (BAPPENAS and BAKORNAS PB, 2006).

As of 1999 and until a few years ago, Indonesia had been marred by vertical and horizontal conflicts, which were marked by the onset of social unrest in several of its regions, including Sambas (West Kalimantan), Maluku, Aceh, Poso and others, and which saw the displacement of almost a million people over 20 provinces. Though the displacement issue has mostly been addressed, the potential for recurrence at any given time remains in the event of social conflict (BAPPENAS and BAKORNAS PB, 2006).

## **3. Problems**

### *3.1. Mitigation*

#### *a. Hazard mapping*

Hazard mapping in Indonesia is still undertaken separately by respective sector. Mapping for volcanic eruption, earthquake, and landslide is done by the General Director of Geology and Mineral Resources, while hazard mapping for floods is executed by the Directorate General of Water Resources. Meanwhile, the State Ministry for Environment has also made a map depicting degrees of forest and land fire and yearly flood proneness every year for the purpose

of preventing and controlling environmental impacts. Then, it is necessary to integrate maps of various types of hazards into an integrated map. Subsequently, the established hazard maps need to be shared with the relevant local communities.

BAPPENAS and BAKORNAS PB (2006) have noticed this problem and put the dissemination of hazard zone mapping and its integration onto the land use mapping as the national policy and strategy on disaster management.

#### *b. Information system*

Several institutions in Indonesia are responsible for developing their own information system in line with their function and mandate. The information system takes the form of websites, bulletins, and other communication media. BAKORNAS PB (2005) states that the problem in developing disaster information systems is that each institution develops its own information system based on its main duties. So disaster is only a small part of the delivered information. Based on this problem, the development of integrated disaster management information systems is a must. BAPPENAS and BAKORNAS PB (2006) also note the importance of information exchange among communities at all levels - experts, practitioners, students, government, managers, household, etc.

#### *c. Education*

The subject of disaster management is not part of the school curricula nationally. BAKORNAS PB (2005) states that only a few schools in disaster prone areas teach the subject of disaster management.

Being an essential part of disaster mitigation, UN-ISDR (2006) states that disaster knowledge should be included in formal education. Moreover, BAKORNAS PB (2005) expects that education on disaster management should be conducted at every level of education starting from elementary to secondary school beginners, and university.

At higher education level, several academic institutions have developed disaster research centres, such as Gajah Mada University, Bandung Institute of Technology, Andalas University, Syiah Kuala University and others. The constraints being faced in conducting research and study include lack of public campaigns, lack of facilities, human and financial resources for research (BAPPENAS and BAKORNAS PB, 2005).

Some of the indigenous knowledge on disaster risk reduction among communities also has to be communicated to other communities. The Government of Indonesia (GoI) and UN (2005) note that some population groups (eg. people living on Simeulue Island) retained ancient memories and sought refuge on higher ground, thereby greatly mitigating the effects of the disaster in terms of the number of people killed in those locations. At the same time, in a few areas (eg. Meulaboh), the army quickly reacted to early indicators and started evacuating people.

Further, BAPPENAS and BAKORNAS PB (2006) have raised some challenging points that need to be addressed by researchers:

- to improve the methods for predictive multi risk assessment and socioeconomic cost benefit analysis of risk reduction actions.
- to strengthen the technical and scientific capacity to develop and apply methodologies, studies and models to assess vulnerabilities to and the impact of geological, weather, water and climate-related hazards.

#### *d. Legislation and policy*

Indonesian national policy and strategy on disaster management (Bakornas PB, 2005), has addressed the fact that legislation in disaster prevention and response is one of the key factor in disaster management. Moreover, GoI and UN (2005) also put an emphasis on national and local legal framework establishment.

GoI and UN (2005) have found that the legal framework for disaster management in Indonesia emerged as somewhat weak, fragmentary and at times duplicative. Although complex command and control structures were put into place in response to this exceptional event, in practice there appeared to be no clear, unique attribution of roles and responsibilities among various components of the public administration. *Ad hoc* decrees and regulations were issued to respond to the emergency, creating structures with uncertain power and resourcing which were sometimes duplicative of what already existed. At the local level, different response mechanisms were in place in different areas and insufficient communication among the areas was noted. As an example, the North Sumatra administration was able to function fully on response, while in Aceh Province, where tremendous losses were reported among the ranks of government officials, the situation remained somewhat confused for a longer time.

Responding to this problem, the Government of Indonesia has acted quickly and as a result some pieces of legislation have been produced. They are as follow:

- Law Number 24 Year 2007, about Disaster Management;
- Presidential Regulation Number 8 Year 2008, about National Disaster Management Board (BNPB);
- Government Regulation Number 21 Year 2008, about Disaster Management Implementation;
- Government Regulation Number 22 Year 2008, about Financial and Management Aid for Disaster;
- Government Regulation Number 23 Year 2008, about the Role of International Institution and NGOs, in Disaster Management;

At local government level, for example, the government of West Sumatra Province has also produced a series of Disaster legislation:

- Regional Regulation of West Sumatra Province Number 5 Year 2007, about Disaster Management;
- Governor of West Sumatra Instruction Number 32 Year 2007, about Standard Operating Procedure on Disaster Management and Evacuee Handling in West Sumatra Province.
- And recently, Disaster Management Plan year 2008-2012 for West Sumatra province, which was produced by the Local Planning Development Board of West Sumatra in association with Disaster Study Center of Andalas University .

The challenge in the future is how to disseminate the legislation to all stakeholders. The other problem being faced is the change of the government system from one which is centralized to a de-centralized system. With such a change, the regional government has full authority to set its own institutional arrangements. Consequently, the unit authorized for disaster management in the region is very much determined by the policy of the respective region. Sometimes the authority is quite big, but there are many cases where the unit is only a small part of an organization unit with very little authority. Strong policy on disaster in every regional government is needed.

### *3.2. Preparedness*

#### *a. Public awareness*

The new paradigm in disaster management is that the responsibility for disaster management no longer lies with the government, but is a shared responsibility of all elements of society (BAPPENAS and BAKORNAS PB, 2006). It means that community also has to be responsible for that. Moreover UN-ISDR (2006) also notes that public awareness is an essential element of preparedness for saving lives and livelihoods. Hadi (2007) admits that there is a lack of understanding in the preparation of disaster preparedness and risk reduction in Indonesia. Community Based Disaster Risk Reduction is probably the best answer to fulfill that program. Unfortunately, in Indonesia the implementation of this method is still very limited and new.

#### *b. Training and simulation*

Training and simulation in facing disaster is also a new thing in Indonesian communities. In Padang city, for example, before the devastating tsunami in 2004, no tsunami evacuation drill took place. IRIN (2007) reported that disaster preparedness training, which included evacuation simulation, started in 2005.

The problem with the simulation is that not all the community gets involved. Most of the participants are students. Hence, it is really a big challenge to involve all the community in it; a new and innovative method has to be created. Training materials on disaster are also not easy to get and numbers of trainers are limited.

*c. Infrastructure and urban planning*

Since disaster is not taken into account when planning national or regional development in most part of Indonesia, it has become obvious that the infrastructure can not support disaster preparedness. For example, in Padang city, when big earthquakes occurs, most of its population will be in a hurry to run to the hilly areas. Every time it has happened, bad traffic jams took place.

IRIN (2007) in an interview with Prayitno, a former UN-OCHA consultant, states "in Padang, roads and bridges leading to higher ground are narrow. If you look at evacuation routes for the Padang area, they don't have the capacity for people to evacuate the 7-10km into the hills to high ground." In fact, after an earthquake on 12 September 2007, there was a three-hour jam of cars trying to get up into the hills.

In response to this problem, the government has promised to create a new evacuation route with a wider road, and also nominate several tall buildings to be used in the event of a tsunami. A study on the evacuation route has also been established this year. Designers working in urban or rural planning, especially of disaster prone areas, have to carry out risk assessments in order to minimize the impact of disaster.

*d. Early warning system*

The 26 December 2004 disaster has focused national and international attention on the lack of early warning and appropriate immediate response mechanisms in the face of a tsunami. While in the case of Indonesia, a minimal amount of time was available to warn communities (less than 15 minutes) between the occurrence of the earthquake and the arrival of the tidal surges, lack of established international and national communication channels to alert populations and locations in harm's way has gathered much attention (BAPPENAS, 2005).

Nearly 4 years after the devastating tsunami, on 11 November 2008, the president of the republic of Indonesia launched the Indonesia Tsunami Early Warning System (InaTEWS). The early warning system can reportedly predict and disseminate news of a possible tsunami within five minutes of an earthquake.

The system, which consists of buoys linked to detectors on the seabed, was developed at a cost of 1.4 trillion rupiahs (\$130 million). Germany contributed 45 million Euros (\$56 million) to the project, while China, Japan, the USA and France contributed equipment and technical assistance (IRIN, 2008). Unfortunately, on the day it launched, Padang Ekspres (2008) reported that there are still some problems with the InaTEWS. People in radius 2km from the siren tower in Padang do not hear the warning. In Pariaman district, the tower was hit by lightning.

*e. Capacity building*

Research which was carried out by the National Institutes of Science of Indonesia (LIPI) (2006) in three cities on the west coast of Sumatera, shows that the disaster capacity index of household, school and government are low. The school capacity index is the lowest among them (Table 2).

**Table 2.** Disaster preparedness capacity index in Aceh Besar, Bengkulu dan Padang (source: LIPI (2006))

Location	Stakeholders		
	Household	School	Government
Aceh Besar	57	50	48
Padang	56	59	75
Bengkulu	51	48	54

Rafliana (2006), LIPI staff, in Koran Indonesia tells us that based on LIPI research on 6 regional area in Indonesia, it was found that the lowest capacity index in facing disaster is that of schools. Hence, it becomes clear that the capacity index for all stakeholders has to be enhanced.

### *3.3. Response*

The 2004 tsunami has produced a number of significant areas that need to be fixed on response and recovery program. Below are some problems that arise during the emergency and recovery stages.

#### **a. Coordination**

The number of NGOs active in disaster management, mainly during disasters, has created its own problem concerning how to organize them. They feel that they are the ones mobilizing assistance and therefore they are the ones to deliver the assistance. There is a reluctance on the part of these NGOs to be coordinated by the government. People participation in disaster management currently is fragmented following what each NGO will do, lacking integration (Bakornas PB, 2005). Therefore, there is a need for regulation that will govern the mobilization and delivery of the resources for assistance to the affected population.

GoI and UN (2005) and EWC (2005) have also noted the coordination problem. The sudden activity of large numbers of aid providers in tsunami-affected areas overwhelmed the capacity of states to effectively coordinate relief efforts. The efforts of multiple institutions and organizations providing relief were not harmonized because of a lack of coordination among humanitarian and aid agencies, different levels of government, competing agendas, and lack of NGO accountability.

On the national level, coordination in disaster responses is done by the government in a graded manner starting from the national level to the provincial level and district/municipality level through the mechanism of Bakornas, Satkorlak, and Satlak. Satlak together with the community at the district/municipality level is responsible for directly handling every disaster occurrence. When the disaster is beyond their coping capacity, Satlak will ask for help of Satkorlak and Bakornas. If the disaster is so big and of national level scale, the disaster response is taken by Bakornas (Bakornas PB, 2005)

#### *b. Discrimination*

East-West Center (EWC) (2005) reports that a number of vulnerable populations, in particular women and members of certain ethnic or religious groups, did not receive equal assistance. Its research revealed multiple causes for maldistribution of aid, including discrimination towards certain ethnic, religious, or marginalized subgroups such as castes; inequities based on political influence; bureaucratic inefficiencies; and exclusion of specific groups based on government definitions of victimhood.

#### *c. Funding*

Emergency funding at the national level is allocated for ministries related to emergency responses such as Health, Social affairs, and Public Works Ministries. The Social Affairs Ministry is mandated to provide food assistance and other social assistance. The Health Ministry is responsible to provide health and medical care services in emergency situations. The Settlement and Regional Infrastructure Ministry is responsible for providing temporary shelters, sanitation, potable water and emergency reconstruction of critically required infrastructure. Bakornas PB (2005) noticed that the problem with emergency funds is its too limited amount when compared to the costs for responding to the emergency situation. In

contrast, the arrangement at the regional level very much depends on the capacity of individual regions.

#### *d. Information management*

As commonly applied in the majority of other post-disaster recovery activities, since the beginning, information management is one of the challenges that cannot be entirely planned. As a result, the implementation of data collection is disintegrated.

BRR (2007) reported that among hundreds of recovery work implementers (NGOs/donors) only a small portion diligently-regularly and timely-update BRR's centralized information system (Recovery Aceh-Nias Database, RAND), whereas actually, the database was designed to help them and monitor the recovery projects. Such disintegration and indolence frequently results in the availability of abundant data which is not appropriate to be used as "information", let alone as the "lessons learnt" with regard to gaps, bottlenecks, overlapping programs and so forth. To this date, there are no systems which cover the needs for management of large-scale projects like those implemented in Aceh and Nias. However, the issue of disintegration in the database does not stem from the system. The main problem lies in the difficulty to integrate the reporting systems of various government or non-government agencies.

Furthermore, BRR (2007) states that Non-uniform output indicators and accountability which is often not fully transparent or different priorities have also posed another problem, in such a way that efforts made to consistently integrate data from all recovery implementers has not been as easy as turning one's hand.

By taking into account the above problems, in the future information management systems and methods both inter government agencies and inter institutions must be handled in a more proper and integrated manner. Responsibility on data consistency, especially those related to the key performance indicators, must be prioritized.

### *3.4. Recovery*

#### *a. Accountability*

BRR (2007) informs us that the issue of the implementers' accountability to the donors also plays a role in the problem of information traffic and uniformity. There are two accountability alternatives that are extant. Accountability can be shown through operational reports and their procedures or through reports on the program absorption by the beneficiaries.

Ideally, those two types of reports are incorporated into one comprehensive report. However, this is not easily realized. The basic and urgent needs of the beneficiaries are often contrary to the procedures and the donors' standards that must be complied to by recovery implementers. Donors expect that their assistance reaches and is properly used by the beneficiaries. For that purpose consistent standard procedures are required. However, viewed from the perspective of urgency and optimal and even distribution, such standard procedures can be a source of frustration. As a result, accountability becomes something that tends to be a matter of standardization rather than optimization.

One of the examples is the housing issue. In one village, there is a large-sized house which has been provided in accordance with European standards, for example. While right beside this house there is a smaller-sized houses provided by another donor with different standards. In Aceh and Nias, about 120,000 houses need to be construction from funds originating from various NGOs/donors with a variety of different standards. In this respect, standardization can cause not only unnecessary delays but also inequity of output. In addition, the time needed for applying different standards can differ greatly, and so too can the progress indicators (BRR, 2007).

Reports from the recovery implementers must certainly be based on a form of standardization, but not at the expense of ignoring optimum output to the beneficiaries. Therefore, a compromise must be sought between standardization and optimization to ensure that the interests of the implementers and the beneficiaries and both fulfilled. Indonesia has standards suitable for the local needs. If this standard is applied, there will be at least a common



solution. Furthermore, the more uniform the standardization of each implementer with the same indicators, the easier it is to collate and process statistical data (BRR, 2007).

*e. Budget estimation*

BRR (2007) states that there are also lessons learnt in the fields of logistics, inflation - which was not sufficiently considered when estimating the budget ceiling - and so forth. As taken from above, a new knowledge of recovery programs and activities can be established.

*c. House ownership and land acquisition*

The problem of house ownership still remains a constraint within the housing programs. And land acquisition still requires the full attention of the National Land Agency (BPN), both at the central and regional levels. The development of settlements also requires an integrated effort from all sectors to ensure a quick and comprehensive fulfillment of the gaps (BRR, 2007).

*d. Lack of community participation*

Government and relief officials often failed to consult survivors and their communities about decisions regarding aid distribution, resettlement, and reconstruction aid. In some cases, these officials discredited or ignored the views and opinions of local communities. Donors and aid agencies often prioritized timely outcomes over deliberative processes that allowed for community participation and discussion (EWC, 2005).

*4. National policy*

As stated earlier, there is a shift of paradigm in Indonesia on facing disaster. Hadi (2007) summarizes the changes which are mandated by the Indonesian Law no 27 of 2007 about Disaster Management (Table 3).

**Table 3.** Shifted paradigm in Indonesia on Disaster (source: Hadi, 2007)

Old Paradigm	New Paradigm
Emergency response	Risk management
Protection as a blessing given by the government	Protection as the people’s human right
Handling disasters as an extraordinary issue	Handling disasters as the daily task of administration and development

Moreover, Indonesia has created a national policy and strategy in disaster as follows (Bakornas PB, 2005):

- To strengthen public awareness and preparedness of local government with the priority on disaster prone areas
- to disseminate hazard zone mapping to the local level and its integration to land use mapping
- to strengthen the capability in disaster detection by providing infrastructure and human resources
- to develop disaster management information system
- to strengthen people capacity through training and education
- to establish integrated rapid and accurate response mechanism
- to issue procedures and guidelines in disaster management, and
- to prepare legislation in disaster prevention and response

In addition, BAPPENAS and BAKORNAS PB (2006) state that there are five key priority areas for disaster risk reduction that must be addressed. The activities to be developed are as follow:

- Ensure that disaster risk reduction is a national and a local priority with a strong institutional basis for implementation.
- Identify, assess and monitor disaster risks and enhance early warning.
- Use knowledge, innovation and education to build a culture of safety and resilience at all level.
- Reduce underlying risk factors.
- Strengthen disaster preparedness for effective response at all levels.

#### 4. Conclusions

There is a new paradigm in Indonesia in managing disaster, from emergency response to disaster risk management. However, some problems are still present in any part of the disaster management cycle. There are problems of stakeholder capacity, management and the information system, legislation and policy, infrastructure, and community participation.

#### References

- Bakornas PB. (2005) *Indonesia: national information*. Country Report. World Conference on Disaster Reduction. Yokohama, 18-22 January 2005.
- Bappenas. (2005) *Indonesia: notes on reconstruction, the December 26, 2004, natural disaster*. The Consultative Group on Indonesia. Jakarta, 19-20 January 2005.
- Bappenas and Bakornas PB. (2006) *Republic of Indonesia: national action plan for disaster reduction 2006-2009*.
- BRR. (2006) *Aceh and Nias, 2 years after tsunami*. 2006 progress report.
- BRR. (2007) *Strengthening the narrative of community life: two years report*. Executing Agency for the Rehabilitation and Reconstruction of Aceh-Nias.
- EWC. (2005) *After the tsunami: human rights of vulnerable populations*. <http://www.reliefweb.int/rw/rwb.nsf/db900sid/VBOL-6JEE3A?openDocument&rc=3&cc=idn>
- GoI and UN. (2005) *Post tsunami lesson learned and best practices workshop*. Jakarta, 16-17 May 2005
- Hadi, S. (2007) *Mainstreaming disaster risk reduction in Indonesia (Stage 2 application for GFDRR track II support)*. Presented on Climate Change, Disaster Risks and Poverty Reduction: Towards a common approach to reduce vulnerability Organized by SIDA. UN-ISDR and the World Bank Stockholm. Sweden, 23-25 October 2007.
- IRIN. (2007) *Indonesia: disaster-preparedness drive in West Sumatra*. <http://www.irinnews.org/Report.aspx?ReportId=76049>
- IRIN. (2008) *Indonesia: Tsunami early warning system launched*. <http://www.irinnews.org/Report.aspx?ReportId=8143>
- JRF. (2007) *One year after the Java earthquake and tsunami: reconstruction achievements and the results of the Java Reconstruction Fund*. Progress report 2007.
- LIPI. (2006) *Tingkat kesiapsiagaan masyarakat di pedesaan kabupaten Aceh Besar, kota Bengkulu dan Padang*. <http://www.siagabencana lipi.go.id/index.php?q=node/18>
- NIA. (2004) *Indonesia 2004: an official handbook*. National Information Agency Republic of Indonesia.
- Padang Ekspres. (2008) *Siren is not heard, president appreciate West Sumatra*. <http://www.padangekspres.co.id/content/view/22743/>

- PreventionWeb. (2008) *Indonesia Data and Statistics*. <http://www.preventionweb.net/english/countries/statistics/?cid=80>
- Rafliana, I. (2006). *Rendah, mitigasi bencana sekolah*.  
<http://www.koranindonesia.com/2007/12/17/rendah-mitigasi-bencana-sekolah/>
- Shaluf, I. M. (2008) *Technological disaster stages and management*. *Disaster Prevention and Management*. Vol. 17 No. 1. pp. 114-126
- TEC. (2006) *South Asia: the Role of needs assessment in the tsunami response*.  
<http://www.reliefweb.int/rw/rwb.nsf/db900sid/YAOI-6S99PU?OpenDocument&rc=3&cc=idn>
- UN-ISDR. (2006) *Lessons for a safer future: drawing on the experience of the Indian ocean tsunami disaster*. UN-ISDR report.
- UN-ISDR. (2008) *Terminology: basic terms of disaster risk reduction*.  
<http://www.unisdr.org/eng/library/lib-terminology-eng-p.htm>.
- USGS. (2008a) *Indonesia political map*.  
<http://earthquake.usgs.gov/regional/world/indonesia/map.php>
- USGS. (2008b) *Historic world earthquake*.  
[http://earthquake.usgs.gov/regional/world/historical\\_country.php#indonesia](http://earthquake.usgs.gov/regional/world/historical_country.php#indonesia)
- USGS. (2008c) *Largest and deadliest earthquakes by year*.  
<http://earthquake.usgs.gov/eqcenter/eqarchives/year/byyear.php>

# The Retrospective Iconicity of '*Guerrillero Heroico*'

Mike Parker, *Department of Journalism and Social Studies, University of Central Lancashire*

## Abstract

Rightly described as the most recognisable photograph in history, the 1960 image of Ernesto 'Che' Guevara by the Cuban photographer Alberto Korda is a universal cultural icon of the highest order. In this paper I will draw on my ongoing research into the iconicity of cultural imagery and argue that, but for the death of Guevara on October 9th 1967, the image would not have become iconic on the scale that it is today. To support this argument I will examine the imagery of the 1964 Russian/Cuban film *Soy Cuba* (I am Cuba) arguing that the prominent imagery of the film focused on stylised revolutionary symbols with only one individual- Fidel Castro- being strongly characterised. Guevara was far from a well-known individual prior to his assassination in 1967; the focus of the paper will extend the idea of Guevara's anonymity and explore the retrospective embellishment of the Guevara myth by civil protest groups, especially those in France, in the months following his death.

## The Retrospective Iconicity of '*Guerrillero Heroico*'

Ernesto 'Che' Guevara died on October 9<sup>th</sup> 1967. Brutally executed by Bolivian soldiers, with the help of the CIA, it was the final act of a life dedicated to permanent revolution in countries perceived, by him, to be dominated by Western capitalist imperialism. 'Terrorist' or 'freedom fighter'- depending on political stance- it is difficult to underestimate the impact that one specific image of Guevara has had upon the international cultural mainstream. In *Che Guevara: Revolutionary and Icon*, Hannah Charlton makes this observation:

Walk through any major metropolis around the globe and it is likely that you will come across an image of Che Guevara...this single image of a face may be the most reproduced image in the history of photography. It has been endlessly mutated, transformed and morphed- and as such tells the history of the last forty years of visual, popular culture (qtd in Ziff, 2006:1)

In this article I will examine the series of events that led, in early 1968, to the formation of a universal 'high order' cultural icon. I will present the argument that the heightened receptivity to the single image of Guevara, Alberto Korda's 1960 *Guerrillero Heroico* photograph, offers a unique opportunity to examine the 'fixing' of symbolic meaning to iconic image. A crucial factor of this argument is that, contrary to popular belief, Guevara was not a well known 'celebrity' figure during the 1960s, and that, from this, his enduring iconic status is a direct result of the retrospective embellishment of his biographical detail creating the narrative of the iconic symbol. Contrary to recent work on the Guevara icon (Larson, Lizardo, 2007) (Ziff, 2006) who see the icon as fluid and malleable to the needs of specific interest groups (in particular, commercial enterprise and radical political organisations), my research suggests that the meaning inherent in the Guevara icon was fixed during a very short period of time in the spring/summer of 1968 and that it is the rigidity of this fixed meaning that has contributed to the overall durability of the iconic form. The first section of the article will present evidence that challenges the common idea of Guevara as a well known public figure prior to his death in 1967. From this I will argue that the personalisation of sentiment expressed by large numbers of people on encountering his image in early 1968 reflects, not a sense of loss for a well-loved and well known individual,<sup>1</sup> but the heightened receptivity (of specific subject groups) to an extraordinary image at exactly the right moment in history. In this respect I will focus on the French civil unrest of 1968 as being crucial to the retrospective sanitising of the Guevarian myth.

---

<sup>1</sup> I particularly wish to oppose the view that the tragic outpouring of public grief was similar in Guevara's example to that of other 'high order' iconic images, such as John F Kennedy, Princess Diana and, recently, Michael Jackson.

Differentiating genuine retained imagery in common memory (cultural icons) from the increasingly hyperbolic use of the term in the popular media is a recognised problem in image studies (Meades, 2009), and to help overcome this I have adopted a working definition of cultural iconicity which, I believe, assists comprehension in this area. Real cultural icons have characteristics that are readily discernible; they are always image based, and the imagery tends to be distinct, durable, and reproducible in common memory.<sup>2</sup> In addition to this I believe that 'high order' cultural icons exhibit a dramatic narrative between the inherent meaning of the image form and the receptive subject. More often than not the dramatic narrative of the icon is heightened by a tragic incident directly related to individual/object concerned. Obvious examples of this would be individuals such as Marilyn Monroe, James Dean and Princess Diana, yet less obvious examples often yield interesting dramatic factors that help explain the durability of the image. The cultural history of the Guevara icon is a singular example of the tragic/dramatic narrative of a 'high order' symbolic image being created in an intensely volatile period.

### **Soy Cuba**

The critically renowned 1964 film *Soy Cuba* (I Am Cuba)<sup>3</sup> by the celebrated Soviet director Mikhail Kalatozov gives us firm evidence that the iconic image of Guevara was largely unknown prior to Guevara's demise. Innovative both technically and visually, the film was poorly received by both sides of the joint Cuban/Russian production: the former criticising its stereotypical portrayal of Cubans while the latter castigated its lack of revolutionary vigour. The complete lack of the Guevara iconic image in the film is strong evidence that, even in his adopted country, Guevara, at this particular time, was not a symbolic individual. Following the lives of several ordinary Cubans the first half of the film deals with the impact of American colonialism and the often devastating effect this had on both the peasant and urban population of Batista's regime. The second part of the film concentrates on the impact of Fidel Castro's disastrous armed landing with 82 guerrillas (including Guevara) and the subsequent regrouping of the 15 or so remaining fighters (again including Guevara) in the Sierra Maestra mountains in Cuba. The film itself is now critically recognised as a masterpiece of its craft, utilising stylistic framing and long study shots to achieve high levels of emotional connection and dramatic effect. Several of these shots are signature shots of Fidel Castro, adorned with beard, cap, cigar and rifle - all associated imagery of the Castro icon. Much of the revolutionary impetus of the film is achieved by dramatic portraiture of the peasant farmer Mariano - an image of a heroic, recalcitrant young man - white, handsome, unkempt, long lank hair blowing in the wind. Even though Alberto Korda's photographic image was available (it was shot 4 years prior to the film) together with the likelihood that Guevara may have been fairly recognisable in his adopted country, this was not a reproduction that was influenced by the *guerrillio heroico* image. Only two extremely brief shots in the film could possibly be seen to represent a signature of the classic Guevara icon (beret, star, and beard) but these are so fleeting that it is overwhelmingly likely that Kalatozov used an established visual idea of the Cuban guerrilla/revolutionary rather than being inspired by Guevara's specific image. The primary reason for this is that, at the time, both Guevara the individual and Guevara the image/icon were unfamiliar in common memory.

First published on April 16<sup>th</sup> 1961 in the Cuban newspaper *Revolución* to advertise a conference featuring guest speaker Dr Ernesto 'Che' Guevara, and again (after the conference was aborted due to the 'Bay of Pigs' invasion) on April 28<sup>th</sup>, the initial publication of Guevara's image had almost no impact (in terms of its iconicity) and, for well over 6 years, there are no other instances of the image being used. During this period Guevara was a member of the newly developing Cuban Marxist government, holding executive positions and operating as one of Fidel Castro's most influential confidants; yet, despite the Cuban situation from the late 1950s being politically high profile, Guevara was simply one of a large number of individuals who formed a backdrop to the larger picture in Cuba. It is, I argue, after the fact (being his

---

<sup>2</sup> For a more detailed analysis of this definition see Parker (2007)

<sup>3</sup> Director: Mikhail Kalatozov. Writers: Enrique Pineda Barnet, Yevgeni Yevtushenko (Cuban version). The film was rediscovered and released in 1995 with the support and guidance of prominent directors Martin Scorsese and Francis Ford Coppola to overwhelming praise for its accomplished, artistic cinematography. See, <http://www.avclub.com/articles/the-new-cult-canon-i-am-cuba,2282/>

death) that our retrospective knowledge of Guevara's iconicity has been focused to include his biographical detail together with retrospective interest in other images of the man during the period in question. My argument then is that Che Guevara may not be known to us today if he had survived capture and assassination attempts in Bolivia in October 1967. The dramatic element is crucial for the correct comprehension of iconic imagery, and there are few better examples in history of the socio/cultural impetus being so centred on one distinct image. The moment of Guevara's death imploded the international ferment of the period onto a single study which, in turn, became a universal reservoir for anti-authoritarian meaning.

## The Guevara Myth

What I would like to do now is to take a close look at the personalisation of the Guevara legend, arguing that the meaning of the icon was retrospectively accommodated in order to synthesise the emotional and mythological demands to a perfect iconic form. The cult of Che - the iconography of the image - has deified the myth of the man to the extent that all of his actions and all his words now seem irreproachable in their moral worth; as Andrew Sinclair suggests, 'he seems to have had hardly any contradictions or inner conflicts. He was amazingly consistent in all he thought and said and did...his consistency was almost maddening in its effortlessness' (Sinclair, 2006: 181). When Guevara died and Korda's 'Christ-like' image became popularly known, the sense of loss was felt personally by large numbers of individuals who, in some shape or form, were fighting battles of their own. It was as though someone's brave big brother had suddenly been taken away, as Stokely Carmichael comments: 'Che Guevara is not dead. I do not wish to speak of Che as if he were dead. It would not make sense...I never met Che in person, but I know him' (Sinclair, 2006: 67) or as Andrew Salkey notes, 'the news of his death, the most humbling and numbing I have experienced, is similar in personal loss to a death in one's own family' (Sinclair, 2006:113).

It is clear that Guevara's death caused a widespread, deeply felt, emotional response. My research suggests, however, that the emotive effect of the image triggered (in the receptive subject) the need for a retrospective embellishment of the Guevara myth to correlate the meaning with the image. Che Guevara was an international insurrectionist who moved from country to country in the middle 1960's to lead peasant revolts against political systems that were perceived by him to be imperialistically corrupt, and to achieve this he needed to be strictly incognito. I have suggested that, even during the high public profile of the 1959 Cuban revolution, Guevara would not have been a particularly well known individual, even in his adopted home, and from 1964 he was fighting undercover in Guatemala, Congo and Bolivia, well out of the public eye. This comment from Peter Kornbluh - senior analyst at the American National Security Archive - is particularly revealing: 'when he resurfaced in Bolivia in 1967 you see the CIA intelligence reports and they say "well I guess we were wrong in thinking that he might have been dead these last several years because it appears that he is alive and well and promoting revolution in the jungles of Bolivia"'.<sup>4</sup> So, although some may have been aware of his existence through his writing,<sup>5</sup> it is likely that Guevara, for the rest of the world and shortly before his death, remained relatively unknown<sup>6</sup>. Then he was captured and killed, and this event generated a unique series of events in the history of image literacy. Perhaps the most significant of these events was the promotion of the 'attractive' characteristics of the Guevara's biography and the disregard of other, more repellent features.

A significant element of the Guevara myth is that he was a selfless combatant for the oppressed masses of the Latin American third world. As an Argentinean he became, as previously stated, a prominent member of Castro's post-revolutionary government and a full

---

<sup>4</sup> Taken from an interview given in the TV documentary *The True Story of Che Guevara*, accessible on <http://www.video.google.com/videoplay 1:08:37/1:30:20>

<sup>5</sup> He was a studious political pamphleteer whose most incendiary work was, possibly, *Guerrilla Warfare* first published in 1960.

<sup>6</sup> The possible exception to this, as shall be explained in greater depth later, was in France where intellectuals such as Sartre, de Beauvoir and Debray became intrigued at Guevara's apparent physical realisation of their existential principles.

Cuban citizen who then opted to fight internationally, 'other nations of the world summon my modest efforts of assistance'<sup>7</sup> It is evident, however, that Guevara was an inveterate traveller who quickly became bored with the repetitiveness of civil life, as his father notes,

The first thing to note is that in my son's veins flowed the blood of the Irish rebels, the Spanish conquistadors and the Argentinean patriots. Evidently Che inherited some of the features of our restless ancestors. There was something in his nature which drew him to distant wanderings, dangerous adventures and new ideas (qtd in Iosif, 1976: 5)

It was all a lot of fun, what with the bombs, speeches, and other distractions to break the monotony I was living in.<sup>8</sup>

Castro's formal explanation for Guevara's renunciation of his Cuban citizenship and subsequent removal from office masks the strained relationship between the Cuban leader and one of his senior ministers.<sup>9</sup> It is also now evident that Guevara was less than successful in his official duties and simply could not adjust to the restrained lifestyle of the government minister. So, rather than Guevara offering his services selflessly to the ideals of revolution for the oppressed masses, it is reasonable to suggest that this was an individual adept at the wandering lifestyle and completely at ease with the rigours of guerrilla warfare (Anderson, 1997).

A more sinister aspect of the mythology is that Guevara was by nature a humanist, merciful soldier who, as a medical doctor, treated his opponents with dignity and care, as in this suggestion by Guevara himself: 'let me say, at the risk of appearing ridiculous, that the true revolutionary is guided by great feelings of love'(Ziff, 2206: 30); or this, from Fidel Castro: 'Che distinguished himself, not only as a fighter, but as a doctor, helping both his wounded comrades and the enemy soldiers' (Sinclair, 2006: 18). The sanctification of Guevara's biography continues in the present with hagiographic depictions in films such as *The Motorcycle Diaries* and *'Che'* (2008). The former example is based on Guevara's own diary of his youthful, 'gap-year', travelling with his friend Alberto Granados. The film is punctuated with Guevara's developing awareness of the perilous conditions in which large numbers of the Latin American peasant class lived, largely as a result (the film suggests) of the corrupting impact of 'Yankee' imperialism. We are given here a stark contrast between good (the downtrodden native class) and bad (dispossessing economic systems) that fire the revolutionary spirit (and with it the high moral stance) of the individual readying for selfless battle against Goliath opposition. The continuation of the mythologizing of Guevara's life through this and many other examples, is, I believe, a necessary extension of the fixing of meaning to image in the months following his death. The tainting of Guevara's character at the crucial point of implementation of the tragic narrative would have dramatically impacted on the potential iconicity of the image. It is clear that the image of the active, ascetic, guerrilla fighter would have been a widely recognisable and, possibly, an attractive symbol in the late 1960's, given the number of worldwide conflicts during the period. Guevara's personal success as part of the small band of individuals who, under Fidel Castro's leadership, overthrew Batista's regime would have proven to be an equally attractive supplement to Korda's image. The severity of revolutionary struggle - the need to kill and to take stern decisions under the most difficult martial conditions - again adds to the allure of both Guevara the man, and his widely disseminating image, especially when this is tempered with his, alleged, humanistic character. The reality of Guevara both as a guerrilla leader and subsequent Marxist despot, however, strain this aspect of the Guevara icon and strengthens the argument that this meaning was a necessary reflection of the needs of the receptive consumers of the image during its initial inception. The iconography

---

<sup>7</sup> Guevara, E. (1965, April). Farewell letter from Che Guevara to Fidel Castro. Retrieved April 9th, 2009, from Marxists.org: <http://www.marxists.org/archive/guevara/1965/04/01.htm>

<sup>8</sup> From a letter Guevara sent to his Mother from Guatemala, cited in Llosa, A. (2005). *The Killing Machine: Che Guevara from Communist Firebrand to Capitalist Brand*. *The New Republic*.

<sup>9</sup> At the time it is well noted that Castro and Guevara favoured different approaches to building Cuba as a communist state, the former being drawn to Soviet interventionism while the latter was drawn to the Chinese Maoist model.

of Guevara – the correlating myth supporting the drama of the image - is formulated around the heroic, selfless, compassionate, brave fighter (whose cause is the elimination of capitalist enslavement) and who never flinched from taking 'severe' 'necessary' actions to further the revolutionary cause. Recent evidence <sup>10</sup> questions the validity of the humanistic element of the Guevara myth; this from Javier Arzuaga chaplain of La Cabana<sup>11</sup> during Guevara's period of command:

There were about eight hundred prisoners in a space fit for no more than three hundred: former Batista military and police personnel, some journalists, a few businessmen and merchants. The revolutionary tribunal was made of militiamen. Che Guevara presided over the appellate court. He never overturned a sentence... I pleaded many times with Che on behalf of prisoners. I remember especially the case of Ariel Lima, a young boy. Che did not budge (qtd in Llosa, 2005: 3).

In addition to his eagerness to exterminate the threat of perceived enemies of the state, Guevara is now accused of other crimes against humanity (Llosa 2005), including his introduction of concentration and forced labour camps, and the imprisonment of 'undesirables' such as homosexuals and drug takers. It is also evident that, rather than finding himself faced with the intolerable dilemmas of leadership in armed conflict, Guevara's well documented statements indicate a fanatical homicidal zeal that developed from being a natural inclination to murder into a genocidal fervour:

I ended the problem with a .32 calibre pistol, in the right side of his brain....his belongings were now mine (Llosa, 2005).

This country is willing to risk everything in an atomic war of unimaginable destructiveness to defend a principle: If the rockets had remained, we would have used them all and directed them against the very heart of the United States, including New York, in our defence against aggression (Llosa, 2005).

I have argued that, as an individual, Guevara could not have been a mainstream figure in public consciousness at the time of his death in 1967. It does not follow from this that the generic image of the Cuban guerrilla fighter was equally unknown. Since 1957 image awareness of the archetypal bearded, fatigue wearing, Cuban combatant would have been increasingly common in Western cultures, especially the USA (DePalma, 2006). There are several thematic aspects of this image that, on the one hand, would have proved attractive to the protest groups we will shortly consider, and, related to this, the guerrilla image may have worked as a preparative image for the specific, detailed meaning of the Guevara icon. As an embodied narrative, the remarkable success of Castro's revolutionary force<sup>12</sup> was a primary symbol of the overthrowing of a powerful opponent - in varying degrees, the USA, colonialism, imperialism, capitalism - by an almost impossibly small cohort. The attractiveness of the Cuban guerrilla symbol correlated closely, I suggest, to the burgeoning romantic rebelliousness of the late 1960s and the concurrent need to capture dramatic meaning within a single image, and nowhere was this need more readily displayed than in the French civil protests of early 1968.

---

<sup>10</sup> For an insight into Guevara's atrocities see James, D. (2001). *Che Guevara*. New York: Cooper Square Press, or, Llosa, A. (2005). *The Killing Machine: Che Guevara from Communist Firebrand to Capitalist Brand*. *The New Republic*.

<sup>11</sup> The Cuban prison governed by Guevara in the first months of 1960.

<sup>12</sup> The actual number of 'revolutionaries' is debatable but of the 160 who attacked the Moncada barracks on July 26th 1953, over 60 were killed in the battle or later after capture, and not all of the remaining individuals joined Castro and his brother on their release from imprisonment in 1955. From this group the initial battle with Batista's Cuban army, following Castro's insurgency of 1959, left only a nucleus of between 12 and 22 guerrillas (Guevara's own estimate was 15) Anderson, J. (1997). *Che Guevara: A Revolutionary Life*. New York: Grove Press. P194. Ramonet, Ignacio, Fidel Castro: My Life. Penguin Books: 2007.



## October 9<sup>th</sup> 1967

The first time I saw the picture by Alberto Korda, I was not even slightly interested in the author. I was only fifteen, and it was the picture that had drawn us - many for the first time - to gather in the streets, crying Che lives! (Ziff, 2006:28)

Guevara's Bolivian diary - the book to which Italian photographer Giorgio Mondolfo is referring to in the above quotation - was printed in 1968 by the wealthy and influential publisher Giangiacomo Feltrinelli. A lifelong communist activist, Feltrinelli visited Cuba prior to Guevara's death and appropriated, for free, the now iconic photograph from Alberto Korda. The inception of the Guevara image as Trisha Ziff suggests, 'seems to have taken place almost immediately after his death' (2006: 19). I argue that an extraordinary coincidental series of events around the time of Guevara's demise, together with the socio/cultural flux of the decade to come, offers compelling insights into the history of embodied cultural iconicity.

Prior to the recent discovery of an article in the French magazine *Paris Match* there was some doubt as to who produced the first official (post 1960) reproduction of Guevara's image. The image was actually used in Cuba shortly after it was taken by Korda in 1960 but then, as previously stated, remained effectively dormant throughout Guevara's life. In the *Paris Match* article entitled '*Les Guerilleros*' the journalist Jean Larteguy asked the poignant question "where is Che Guevara?" Ziff, 2006: 19) next to the now iconic Guevara image. Although Ziff correctly suggests that, '*Paris Match* was clearly responsible for the image being seen and disseminated for the first time not only in France, but also to a wide European audience' (2006:20), my thesis suggests that the primary catalyst for the cultural integration of the icon was the death of Guevara himself. So, while the image was in some form of currency via the *Paris Match* article, it was not utilised as a full embodied icon until several weeks later, fully supporting the idea that the iconic event requires the kind of tragic occurrence epitomised by Guevara's death.

The months following October 9<sup>th</sup> 1967 saw a proliferation of the Guevara image in cultures where the receptive conditions for the myth/image symbiosis could be most effectively met. On October 18<sup>th</sup> 1967 Fidel Castro gave a memorial service in which he eulogised the bravery and revolutionary fervour of Guevara against a backdrop of Korda's photograph reworked into poster form by the Cuban artist José Gómez Fresquet who, sensing the urgency of the moment, worked through the night to produce a graphic reproduction which filled the side of a five story building (Ziff, 2006: 20). Today the building has a steel facsimile of the original poster to stand as a permanent reminder (as if one were necessary) of the power of the icon witnessed by over one million Cubans on that particular day.

To gain an idea of the dramatic narrative synchronised onto the image at the time of Guevara's death consider this extract from Castro's memorial speech:

But those who boast of victory are mistaken. They are mistaken when they think that his death is the end of his ideas, the end of his tactics, the end of his guerrilla tactics, the end of his theory. For the person who fell, as a mortal person, as a person who faced bullets time and again, as a soldier, as a leader, was a thousand times more able than those who killed him by a stroke of luck. (qtd in Ziff, 2006:71)<sup>13</sup>

Leaving aside the animosity that had developed between Castro and Guevara - during his capture in Bolivia Guevara repeatedly told his captors 'Fidel betrayed me' (de Guzman, 1998) - I suggest that the seminal instances of iconic reproductions of the Korda/Guevara image, such as the one above, mark a retrospective reconstruction of biography compelled by the dramatic gravity of the image: in a sense it can reasonably be claimed that there was no Guevara - as we know him today - before the impact of the icon. I have previously shown that during

---

<sup>13</sup> Also, Peter Kornbluh, senior analyst of the American Security Archive, suggests that Castro was forced to become an apologist for Guevara in order to dispel rumours that Castro was behind the assassination, *The True Story of Che Guevara*. (n.d.). Retrieved March 25th, 2009, from <http://video.google.com/videoplay?docid=5762714709014580290>

Guevara's most active years he would not have been distinct from a number of possible template Cuban 'revolutionary' images and that his 'ideas', 'tactics' or 'theory' would have been unknown other than to a handful of (mostly radical left-wing) intellectuals. During a very short period of time following his death, however, Guevara - or as I suggest the dramatic narrative myth of Guevara - became internationally famous. The opening quotation to this section shows little interest in the individual underscoring the product and a high interest in the image itself; this probably underestimates the speed in which the Guevara myth/image synthesised after his death. What I would like to do now is to explore the momentum of one, seminally important, instance of the mythologizing of the Guevara image.

## Paris 1968

I argue that the meaning of the Guevara image was fixed and sanitised in the short period of time following his death, and that the subsequent utilisation of the image, as a cultural icon, is as a symbol with fixed meaning yet alterable significance- according to the specific needs of the receptive subject/group. This idea correlates closely to the empirical research of Larson and Lizardo (2007) who argue that, 'when a memory becomes canonical and the basis for other cultural formations, its malleability decreases' (Larson & Lizardo, 2007:17). The period from 1968-1976 was a time of political and social unrest in much of Western Europe and the USA, and Guevara's image was the perfect resource through which to channel the frustrations rooted in dominant ideological hegemonies. My thesis of the embodied, distinct iconic symbol, in this sense, again finds support from Larson and Lizardo who suggest that 'the memory of Che Guevara corresponds with the recent advance of so-called "identities of resistance" (Castells, 2004) that oppose dominant definitions of citizenship and community in favor of sub-national identities' (Larson & Lizardo, 2007: 24-5), but differs in the approach to the relationship between receptive subject and the iconic image. Whereas Larson and Lizardo argue that 'the memory of Che Guevara has been shaped by shifting values, identities, and political orientations that crosscut generational divisions' (Larson & Lizardo, 2007: 26), my interpretation is that the myth/image synchronisation was complete in the early years, if not months, following his death and that it is the *receptivity* of the subject that has proved variable in the development of the icon.

1968 was a momentous year for civil rights movements internationally. In France and Italy escalating student protest resulted in nationwide strikes and political crisis. The French example, in particular, reveals the extent to which a dramatic narrative relationship can be seen to develop during periods of heightened receptive awareness. In comparison to the civil rights movement in North America, which were fuelled by real political and material inequalities, French protest in the period (often referred to as May '68) was instigated far more by intellectual and cultural factors. Early in 1968 the director of Cinematheque Francaise<sup>14</sup>, the eccentric yet popular Henri Langois, was dismissed from his post as curator sparking a wave of dissent from the artistic community both within France and from abroad. On 14<sup>th</sup> February 3000 people, again mostly from the artistic community, gathered to protest against perceived dictatorial action on the part of the French government; what ensued illustrates the character of civil French protest in the months to come. Holding banners bearing slogans such as 'Films Not Cops' the crowd, including many notable celebrity figures from the film world such as Francois Truffaut and Jean Luc Goddard were ruthlessly attacked by French riot police in full combat clothing, and armed with tear gas and riot batons.

The severity of the police reaction confirmed the sense of authoritarianism displayed by the government of the time, especially towards the artistic community and the intellectual 'new left', as Professor Robin Blackburn notes: 'De Gaulle's apparatus of repression in academia was probably stronger in France than anywhere'.<sup>15</sup> On the 22<sup>nd</sup> March 150 students, artists, and left wing groups occupied an administration building at Nanterre University (a satellite campus

---

<sup>14</sup> Paris based state sponsored film theatre and museum founded by Langlois, Georges Franju and Jean Mitry in 1936.

<sup>15</sup> Taken from a video interview with Professor Robin Blackburn, <http://video.google.co.uk/videosearch?hl=en&q=may+1968+100309>

built to facilitate overspill from the University of Paris) the first in a series of events eventually leading to the collapse of the Pompidou government and the near collapse of French society as a whole. Following months of violent clashes between protesters and the police, student groups occupied the Sorbonne University in May which, following its invasion by armed police, became the catalyst for other groups - workers, high school students, and university lecturers - to react against the severity of police action. By the 16<sup>th</sup> May - with President De Gaulle in enforced exile in Germany and over two thirds of the country on strike - the French nation was on the brink of socialist revolution, with the essence of the revolutionary fervour being undeniably youthful, left wing and anarchistic. Listed below are some examples of the graffiti utilised at the height of the 1968 French civil unrest.

Don't beg for the right to live — take it!

The liberation of humanity is all or nothing!

Those who make revolutions half way only dig their own graves!

Run, comrade, the old world is behind you!

We will ask nothing. We will demand nothing. We will take, occupy!

The revolution doesn't belong to the committees, it's yours!

If we have to resort to force, don't sit on the fence!

No to coat-and-tie revolution! <sup>16</sup>

The intense political atmosphere of Paris in the early summer of 1968 was expressed not only through the medium of graffiti, but also through the newly developing technology of mass poster production. I argue that the expression of highly charged socio/political sentiment through this type of medium, in addition to the anti-systemic polarity of the participants, was a crucial preparative stage in the perceptual receptivity to a charged symbol such as the Guevara image. What I would like to do now is to look at some of the symbolism of the period and to suggest that one image, possessing the capacity to reflect the political passion of the moment, would have proved irresistible to the receptive subjects involved.

The amount of posters produced during the French conflict of 1968 was immense (Corkran, 2005) and as David Singer argues was an extremely important avenue of anti-systemic expression: 'the political vacuum precipitated a political void, and this in turn was filled by posters and pamphlets' (Singer, 2002: 19). Although the slogans and posters of May '68' are symbolic of the general fervour of radical left-wing sentiment they also reveal the desire to express these sympathies through an embodied medium: the slogans, for instance were generally disseminated as physical graffiti and the poster art boldly situated in public areas for maximum effect. The expression of confrontation through this medium was a deeply felt personalisation that consciously avoided - as part of a general Marxist ideology - traditional bourgeois methods of artistic production; the messages of the protest were communicated 'in announcements over microphones, painting on walls, announcements in the cinema during the film, texts written on posters, whenever you empty your glass at the bistro, before making love, after making love, in an elevator'.<sup>17</sup> Turning to the posters themselves<sup>18</sup> Christine Corkran (2005) shows the diversity of production at the Atelier Populaire (People's Studio) within the Ecoles Des Beaux-Arts between May and June 1968, a short yet extremely concentrated period of artistic fabrication. Corkran groups the posters into several categories reflecting the varying degrees of effectiveness and poster numbers. The first group of posters reflected the students desire to free the French university system of its 'old world' paternalistic approach to teaching where the students perceived themselves to be mere receptacles of established and outdated knowledge. Partly this dissatisfaction may have been the result of

---

<sup>16</sup> <http://www.bopsecrets.org/CF/graffiti.htm> 120309

<sup>17</sup> Fisera, V. C. (1979). *Writing on the Wall, May 1968: A Documentary Anthology*. New York: St. Martins Press, p125, quoted in Corkran (2005), p63.

<sup>18</sup> The posters can be viewed at various websites including [www.art-for-a-change.com/Paris/paris](http://www.art-for-a-change.com/Paris/paris).

the relatively large numbers of students within the education system (at the time French policy was to offer a place to any individual who possessed the basic entry requirements) and the disparity of opportunity between the results of a university education and the possibility of a worthwhile career. One Poster clearly represents the sentiments of the protesters towards the establishment of a "Universite Populaire" (Peoples University) assisted by worker unification (the spanner in the clenched fist), which connects closely to another poster, worker-student alliance. The ideas underscoring this particular theme reflect the attempt to unify the protesting students and workers in one 'struggle' Also included within student/worker poster movement were sentiments reflecting co-operation between students and immigrant workers and between students and farmers, the latter being a thinly disguised attack on De-Gaulle's paternalistic, hyper-conservative power base (Corkran, 2005:71). Many of the posters utilised in the May '68' protest were targeted directly at the conditions of the French working class, particularly that of the industrial workers. Again this shows a strong desire for unification, in this case unification between workers in different industries and workers of differing ethnic origins with accompanying messages such as, 'Taxi drivers, the struggle continues'.

This heightened sense of protest, and its expression through embodied symbolism, created the perfect conditions in which ideas relating to protest, confrontation and rebellion could be assimilated into one symbolic form. It can clearly be argued that, even though the Korda's photograph was utilised in Cuba in October 1967, and *Paris Match* shortly after that, the mythologizing of the Guevarian legend - the synthesising of meaning to image - was to a significant extent a French phenomenon, with parallel, but less dramatic, instances occurring at around the same time in Italy and other Western European countries. One high profile intellectual supporter of the May '68' student protest was Jean-Paul Sartre, and in a radio interview at the height of the conflict Sartre gave this opinion: 'these youngsters don't want the future of their fathers—our future—a future which has proved we were cowardly, worn out, weary, stupefied by total obedience...Violence is the one thing that remains, whatever the regime, for the students...The only relationship they can have to this university is to smash it' (qtd in Corkran, 2005). An advocate of Marxist ideology and an initiator of atheistic existential philosophy Sartre discovered the embodiment of his thoughts in the living form of Ernesto 'Che' Guevara. Believing action to be the true purpose of life, together with the need to be authentic to oneself, Guevara - or perhaps more accurately the elements in Guevara's life that correlated to Sartre's thought - symbolised all the elements of the underlying ideology of the student protests of 1968. In 1960 Sartre, along with Simone de Beauvoir, travelled to Cuba to witness first-hand the aftermath of the Castro revolution and to interview prominent individuals including Guevara.

All three individuals, Guevara, Sartre, and de-Beauvoir, were on the same platform on March 5<sup>th</sup> 1960 when Alberto Korda took the now iconic photograph (Ziff, 2006: 25). Following Guevara's death Sartre famously referred to him as 'not only an intellectual but also the most complete human being of our age'. Wholly enamoured with the perception of Guevara as a man of action, Sartre contributed significantly to the retrospective sanitising of the myth at precisely the instance that it was required, i.e. when the poster representations of Guevara were gaining rapid popularity. In addition to Sartre and de Beauvoir's lionisation of Guevara, the French activist and thinker Regis Debray was also heavily involved in the myth-making surrounding Guevara's adventures. As a professor of philosophy in Havana during the early years of the Castro's regime, Debray was closely associated with the ideals and methods of Marxist inspired guerrilla discipline. During Guevara's ill-fated Bolivian adventure Debray travelled as a 'journalist' to monitor first-hand the efforts of the small band of revolutionary insurrectionists, the extract below taken from Debray's account of the last days of the group, fully encapsulates his emotional appreciation of Guevara as the archetypal mythical rebel:

And amidst all that, the solitude of Che himself in the middle of the encampment, sitting upright on his stool in the evening, by the fire, his khaki felt hat pushed back, drawing slowly on his pipe, and reading perhaps Leon Felipe's poem, *La Gran Aventura*. There he sat, impassable, absent, barely noticing the comings and goings, the bustle, the clatter of cooking pans, the background of swearing and arguments all round him. He remained motionless and distant, like a man transported beyond this world, letting his eyes rest on the flames from time to time with a slight smile, ironic and confident, as though seeing his own inner monologue, as though savouring in advance the unknown satisfaction waiting beyond the horizon in a profusion of glorious struggles. (Debray, 1976: 143)

Imprisoned shortly before Guevara's assassination, the example of Debray, an active participant in armed revolutionary struggle, could only have added to the potential receptivity of the Guevara icon in France during the period under consideration. My argument is that, at this particular historical moment, the conditions for the creation of the perfect cultural icon were completely in place and that, from this, the meaning of the Guevara myth/image was retrospectively created to match the needs of the receptive subjects.

## The Complete Icon

In the later months of 1968 the enraged students of France, together with groups in other parts of Europe finally found an image in which to invest their emotional hostility, and they embraced it with vigour. In April 1968, as David Singer notes, 'they were finally given a hall for political meetings - which they promptly re-christened for Che Guevara' (Singer, 2002 64). Very soon after this the idea of Che Guevara would be inextricably connected to one image, which, in turn, would prove to be one of the most distinct, durable and reproducible images in common memory. My analysis of the Guevara icon has attempted to present reasons why this image has proved to be universally transferable as a symbol unrestricted by cultural or temporal constraints. People search for images to reflect ideas and feelings, cultural icons - especially ones of high order - seem to encapsulate definite perceptual needs and very few achieve this better than the Guevara icon. To accomplish this the dramatic narrative of the distinct image should, under the terms of my hypothesis, be readily discernible and triggered by a tragic event. Having found a single image in which to discharge the complexity of their feelings, European protest groups in 1968, especially those in France, consumed the Guevara image with fervour: in doing this they simultaneously assigned meaning that correlated to their own heightened socio/political sensibilities. It is now difficult to dissociate the retrospectively embellished Guevara myth from the image itself, proving the intense dramatic narrative between icon and subject. The civil protests in France in 1968, I suggest, were so vehemently anti-systemic that the Guevara image became the perfect receptacle for complex anti-establishment feeling previously expressed through graffiti and poster production. Physically, at the time of his capture in Bolivia, Guevara bore little resemblance to the 'heroic', 'Christ-like' image that many people around the world are now familiar with. As I have indicated, the debate now surrounding Guevara fundamentally questions many of the elements of the meaning underpinning the icon. At the moment of his death, however, none of this seemed to matter. The heightened receptivity of the protest movement to an image such as Guevara's proved too strong for anything other than the complete, retrospective, sanitising of his biographical detail. In the case of Cuba conscious knowledge of the proximity of the island to the mainland States, together with the anti-American inspired success of Castro's insurgents, contributed to the high degree of receptivity which I believe existed in the subconscious minds of the late 1960's student groups and civil rights protesters. Certainly there were elements to the Cuban revolution that excited the sympathetic interest of several left-wing intellectuals of the period, including Sartre, Debray and Feltrinelli. The idea that a small, mostly male<sup>19</sup> band of Guerrillas could take action against the might of American colonialism would have proved an extremely attractive 'real-life' embodiment of their anti-authoritarian, existential theories. I have argued that the creation of high order iconic forms, i.e. ones with distinct, durable and reproducible imagery, are very often personalised through the heightened receptivity of individuals or, through common memory, larger cultural groups. The socio/political atmosphere at the time of Guevara's death in October 1967 focused an enormous amount of emotional attention onto a single image - Alberto Korda's photograph '*Guerrillero Heroico*'. This attention, I argue, was primarily a European, intellectual affair which, nevertheless, was personally felt by very large numbers of people. The product

---

<sup>19</sup> Although females played a role in the Cuban revolution (Guevara's wife can be seen in combat clothing in a photograph from the period) they were often assigned predictably stereotypical roles, cooking, teaching, cleaning etc. It may also be significant that the role of women in the French protests of 1968 was similarly underplayed, "for the most part women performed second-class duties such as obtaining food, cleaning, and taking notes" (Corkran, 2005: p39)

resulting from this series of events was a distinct icon fully equipped with a perfectly synchronised dramatic narrative. From this, the iconography of the image - the rigid meaning inherent within the phenomenon - allowed the icon to be utilised in a variety of forms, by groups requiring a strong political image, by non-Marxist civil rights activists, and by individuals who, to this day, regard the image as universally 'cool'. The utilisation of the icon since 1968, together with the concurrent degree to which the significance of the symbol has altered, is a natural extension of this article.

## References

- Anderson, J. (1997). *Che Guevara: A Revolutionary Life*. New York: Grove Press.
- Blackburn, P. R. (2009, March 27th). Paris Riots 1968. (h. 100309, Interviewer)
- Corkran, C. (2005). Defining Political Action: Posters and Graffiti from Paris 1968. Submitted Thesis, Franklin and Marshall College.
- De Guzman, N. (1998) Interview, *USA Today*, 01.05.98
- Debray, R. (1976). *Che's Guerilla War*. Penguin Books.
- DePalma, A. (2006). *The Man Who Invented Fidel: Castro, Cuba, and Herbert L. Matthews of The New York Times*. New York: Perseus Publishing.
- Fisera, V. C. (1979). *Writing on the Wall, May 1968: A Documentary Anthology*. New York: St. Martins Press.
- Guevara, E. (1965, April). *Farewell letter from Che Guevara to Fidel Castro*. Retrieved april 9th, 2009, from Marxists.org: <http://www.marxists.org/archive/guevara/1965/04/01.htm>
- Guevara, E. (1996, January 22nd). Letters by Che Guevara. *The Militant*, p. Vol. 10.
- James, D. (2001). *Che Guevara*. New York: Cooper Square Press.
- Kalatozov, M. (1964). Retrieved April 2nd, 2009, from FilmIreland: [www.filmireland.net/109/cubaspotemkin.htm](http://www.filmireland.net/109/cubaspotemkin.htm)
- Kunzle, D. (1997). *Che Guevara, Icon, Myth and Message*. Los Angeles: Fowler.
- Larson, Jeff A.; Lizardo, Omar. (2007). "Generations, Identities, and the Collective Memory of Che Guevara", *Sociological Forum*, Vol 22, 22 (4), pp. 425-451.
- Lavretsky, Iosif, translated by A.B.Eklof. (1976). *Ernesto Che Guevara*. Moscow: Progress.
- Llosa, A. (2005). The Killing Machine: Che Guevara from Communist Firebrand to Capitalist Brand. *The New Republic*, <http://www.tnr.com/>.
- Meades, J. (2009) Icon: The Adjective of the Age. *More Intelligent Life*. <http://www.moreintelligentlife.com/story/adjective-age>.
- Morgan John, Welton Peter. (1992). *See What I Mean*. London: Edward Arnold.
- Perez, J. L. (1995). *Essays on Cuban History: Historiography and Research*. Gainesville: University Press of Florida.
- Ramonet, I. (2007). *Fidel Castro: My Life*. Penguin Books.
- Sinclair, A. (2006). *Viva Che*. Gloucestershire: Sutton Publishing.
- Singer, D. (2002). *Prelude to Revolution*. Cambridge, MA.: 2002.
- Staten, C. (2005). *History of Cuba*. Gordonsville, VA, USA.: Palgrave Macmillan.
- Worthington, C. F. (2001). *Wiki/History of Cuba*, Wikipedia: [http://en.wikipedia.org/wiki/History\\_of\\_Cuba](http://en.wikipedia.org/wiki/History_of_Cuba).
- Ziff, T. (2006). *Che Guevara Revolutionary & Icon*. London: V&A Publications.

## Film

- Soy Cuba! (I am Cuba)* [Motion Picture]. Director: Mikhail Kalatozov. Writers: Enrique Pineda Barnet, Y. Y. (Director). (1964).
- The True Story of Che Guevara*. (n.d.). Retrieved March 25th, 2009, from <http://video.google.com/videoplay?docid=5762714709014580290>

# Implications of the initial overt decisions on lung nodule detection from chest radiographs

**Mariusz W Pietrzyk**, *School of Medical Imaging Sciences, University of Cumbria*

## Abstract

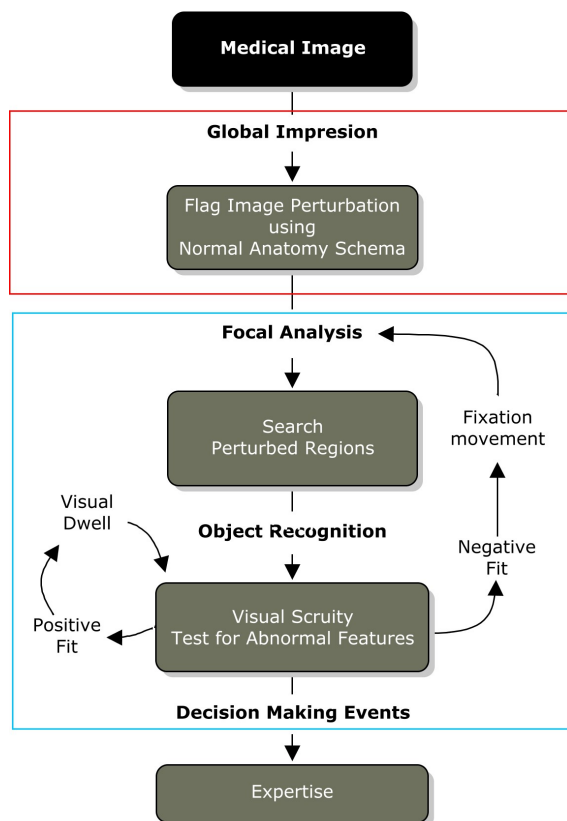
This study focuses on human error in lung nodule detection from a set of posterior-anterior chest radiographs, where combined eye tracking and image analysis methodologies are used. The non-uniform distribution of visual attention spread across the image plane is found with significant differences ( $p < .05$ ) between subjects with different radiological experience. Experts perform better in finding True-Positive (TP) from all the most dwelled True-Negative (TN) regions, avoiding False-Positive (FP). In consequence, the more experienced subject performs with higher accuracy. Simultaneously, the spatial frequency analysis through Wavelet Packet Transforms (WPT) of selected regions of interest (ROI) gives an insight into what kind of information is perceived and processed through neurological pathways during image reading. The significant differences ( $p < .05$ ) between level of experience demonstrate that experts dwelled longer and more accurately than less experienced observers on those locations where local backgrounds are more similar to the nodule containing regions. Furthermore, the probability of the error event in function of the previous outcome correctness is studied. The TP is more likely to happen if it is followed by the correctly recognised and detected nodule-containing region than if followed by incorrectly marked normal ROI. By analogy, the FP occurs more frequently after FP than after TP.

## 1. Introduction

The holistic view generated during the very first hundred milliseconds of perceptual onset seems to have a great influence on the way a particular visual task is performed by human observers (Mello-Thoms 2009). That rapid phase of image perception occurs without eye movements and the visual signal is processed and interpreted through the subject's neurological pathways. That rapid daft interpretation seems to determine where to look next and may be largely responsible for the subsequent visual attention distribution. In consequence, the crucial information for the final decision may be missed or overestimated. Therefore, the misinterpretation of information contained in medical images, such as chest radiographs, is a major source of error in radiology. Information gathering in this context develops through visual searching for a particular target (i.e. lung nodules) in a very noisy background. Search errors occur when observers miss the target without paying any visual attention on it (Manning et al., 2004), while observed, but not reported target events are related to the recognition or detection of that abnormality. Such distinctions are made depending on the value of the cumulative dwell time spent on a particular target. Moreover, recognition or detection errors also happen in cases of reporting normal regions as a target. This difficult radiological task is based on visual perception and multi-choice decision-making processes. Observers not only receive information, but rather, actively develop their understanding of the information through interaction with displayed images. The kind of exploration of visual information in a particular context (an individual patient's case), involves many different factors (Manning et al., 2005). Skill development through years of

practice, training, and experience gathering perceptual skills and knowledge related to the task, enables radiologists to cope with a great deal of information displayed at the image. In consequence, the experts perform diagnostically better than novices, score higher accuracy using fewer gaze points, with shorter dwell times, cover less of the image area compared with the less experienced observers. Those performance parameters have been measured with eye tracking devices, which record the eye-movements from a particular view distances and provides the eye fixation locations at the invested image (Manning et al., 2004). Nonetheless, radiology suffers in accuracy mainly from cognitive decision-making errors. Thus, a common error – missed lung nodule – occurs in 30% of all chest radiographs investigating lung cancer (Carmody et al., 1981), where 60-70% of errors are due to wrong judgment concerning the well localized abnormality. About 20% of missed nodules have not been fixated at all (Kundel et al., 1990).

Studies of radiologists’ perceptual and cognitive processes during medical image reading involve theoretical modelling of visual perception and searching. The background concept of visual perception as a decision-centred process was founded by Gregory and Rock over twenty years ago. More recently, the descriptive model of radiological image interpretation (Nodine et.al., 1987; Kundel et al., 2004) has been developed into an analytical search model (Chakraborty, 2006), which considers the observation that visual attention is not assigned to all locations in the image equally. The model is based on eye-position recording made by radiologists and introduces the term of decision site, which corresponds to regions examined individually by using foveal vision.



**Figure 1.** Flow chart of a visual searching and perception model applied for medical image perception: upon image onset to the subject, a *global impresion* of the medical image is formed often without eye movements.

In this very short (limited by the time of first



saccade, ~ 100msec) phase, the potential candidates for further detail inspection are flagged. These areas are going to be analyzed via foveation process during the next phase – *discovering scanning* or *focal analysis*. The second phase is much longer and can take from a few seconds (for more experienced subjects) to even minutes (for less experienced readers). The eye movements are recoded during the second phase, in which the subject guides their high resolution visual beam to a particular location. The *global impression* is formed based on low spatial frequency components from the whole visual angle (peripheral vision), whereas, the *focal analysis* of the selected region of interest (ROI) is limited by the central (foveal) vision up to few degrees (2-5) of the visual angle. However, the high spatial frequency components are taken under consideration during foveation process, which occurs upon fixations (dwelled over 100ms).

A diagnostic procedure may be improved by increasing the observers' ability to perceive a particular type of abnormality from the image (Manning et al., 2006). It seems likely that some local properties of the images, such as relative localisation, spatial frequencies in different orientations, contrast at various scales, amplitude and phase of the discontinuity, have an influence on observer performance, guiding their searching and sampling strategy during image reading. This study focuses on human error in the lung nodule detection from a set of posterior-anterior chest radiographs, where combined **eye tracking** and **image analysis methodologies** are used according to a perceptual model (Kundel et al., 2004). The first goal is to determine the **spatial frequency (SF)** characteristics of image locations and comparisons of different overt and covert types of observers' decisions and find out if there are any similarities within different observers' groups: experts (clinicians), two levels of novices (students) and subjects without any relevant experience (naives). Furthermore, the **probability of the error** event as a function of the correctness of the previous decision is studied.

## 2. Methodology

The eye tracking data allows analysis of the allocation of the visual attention spread through the image plane during medical image reading. The retrospective design of the experiment allows comparison of each subject's performance with the Ideal Observer response. The assessment of the accuracy has been done according to the most appropriate statistical method for free-response observer performance (Chakraborty, 2008). The relative distances between the overt decisions location and ground truth were used for classification of the subject's outcomes. Moreover, the time order of decisions is also analysed in terms of the correctness of response following certain prior outcomes. The physical properties of the local background around decision-point sites were investigated in terms of spatial frequency properties. The physical image-based features are generated through a wavelet decomposition of the selected regions of interest centred on the found location coordinates.

### 2.1. Materials

Images databank: The test bank of the chest radiographs contains the 20 digitised a gray scale (8-bits) images, with 50% of prevalence of films with highly salient or camouflaged nodules which are histopathologically proven or simulated. Four computer synthesized nodules are not distinguishable from natural lesions (Ethell 2003). 5-30 mm in diameter nodules were located in the lung field or mediastinum. 23 nodules in total have been spread in 10 cases as follow: 5 singles, 1 double, 2 triples, and 2 quinque cases.

Participants: 40 observers with different radiological background have taken part in the eye tracking experiment. The participants have been recruited according to the level of

experience. Experts, the most experienced professional observer: 8 radiologists and 2 reporting radiographs: age range 35 – 55y (mean 47.2). Next, two levels of novices recruited from 3<sup>rd</sup> and 2<sup>nd</sup> year of Radiography BSc at University of Cumbria. First students' group (age range 20-54, mean 29.4) would have 28 weeks clinical experience, whereas second group (age range 19-48, mean 31.6) at least 12 weeks. The less experienced, naïve observers, come from fields not related with radiological task disciplines (age range 22-40, mean 29.2). Subjects were informed that cases which they are going to investigate may contain single, mutilate or none at all nodules. However participants were not made aware of the prevalence of films with nodules, which should not significantly influence readers' accuracy in detection (Gur et al. 2003).

Workstation: The images were displayed at a Medion MD6454AP LCD Monitor. The observer eye-movements were recording during whole session of image test bank reading, using a Tobii x50 system and ClearView analysis software with 50Hz sampling frequency. The relative difference of NIR light reflections at the cornea are used to estimate the spatial location in the camera field of view (FOV), which is 21/16/20cm (weigh/high/depth) at 60cm from the camera. The accuracy of the gaze locations for both eyes has been estimated to 0.5 degree of visual angle, without drift-effects or compensations errors from larger head movements. Any drift-effects are removed by implementation of binocular averaging.

## **2.2. Experimental design**

The eye tracking experiments were undertaken in a dedicated laboratory, in which ambient lighting could be controlled and optimum viewing conditions set up. Luminance can have an effect on image perception (Brennan et al. 2007). A blackout blind was used to reduce any variability in luminance, and the lighting levels were reduced to 30 lux, as determined by a photometer (Unfors Xi Light detector). Quality checks were undertaken on the monitor using the Society of Motion Picture and Television Engineers test pattern. Subjects were told that the images may be normal or have one nodule or multiple nodules up to a maximum of five on any one image. Their task was to decide on the presence or absence of pulmonary nodules and disregard any other radiological findings. All subjects were shown two chest images, a normal one and one with multiple lung nodules prior to the start of the study. A 9 point calibration procedure was carried out and stored. The databank of radiological cases were analysed individually per radiograph as follow: free search was allowed and subjects were instructed to click using a mouse on each nodule while the image was displayed on the screen. The subject terminated a search by pressing the space bar, which led to the next case, until the last image in the data set.

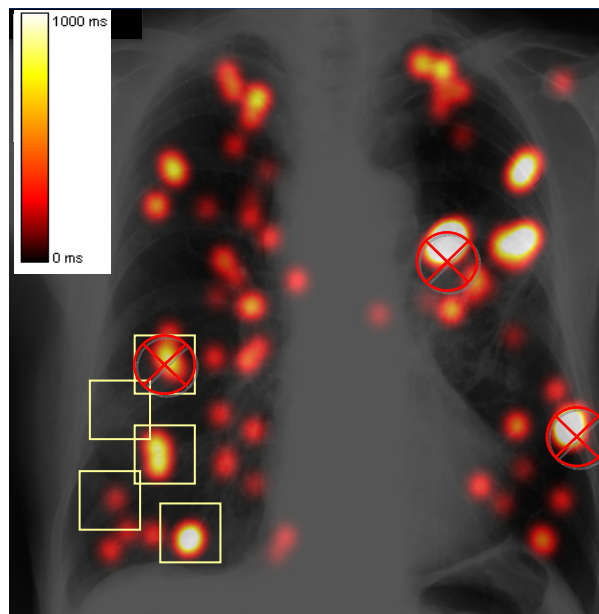
## **2.3. Eye tracking data**

An observer performs a visual investigation by guiding their attention by eye movements to a set of different locations in the image. The reposition of the fovea results from interference of five basic types of movement (Robinson, 1968). Saccades range in duration from 10 to 100ms and occur during repositioning the fovea to a new location. Those rapid movements are too short to allow any visual information to be processed, so the observer stays blind to data capture during these transitions (Shebilske et al., 1983). In acquiring, processing and analyzing information during image reading, the visual performance of a normal observer is maximized at the point of fixation that stabilizes the central retina over an object of interest. So most of the viewing time in visual searching, detection and decision-making tasks is devoted to fixations (Irwin 1992). In summary, eye movement is synchronously shifted with

spatial-visual attention (Hoffman, 1998), and also whatever the type of visual task, saccades and fixation models gain insight into the overt localization of visual attention.

$$F_i(x_i, y_i, t_i, d_i) = \text{cluster}\{ \Sigma_k \text{gaze}_k(x_k, y_k, t_k, d_k) \} : \begin{cases} \forall \text{gaze}_k(x_k, y_k) \in C\{(x_i, y_i), r_i\}; \\ t_i = \Sigma_k t_k(x_i, y_i); \\ r_i = 2 d_i \tan(\theta/2); \end{cases} \quad (1)$$

where  $d_i$  is the view distance measured between the observer and the screen. Thus, a major concern is to reveal the organization of the distribution pattern into sensible groups of fixations. The applied search model (Chakraborty, 2006) assumes an unlimited number of sites, which are calculated through spatially clustering process of fixations with a total dwell time exceeding a certain threshold value. This value indicates that the observer dwelled long enough on the location to make a decision and ranges from an estimated *200ms* (Manning et al., 2004), *330ms* (Chakraborty, 2006; Mello-Thoms et al., 2004; Mello-Thoms, 2006) to *1 sec* (Mello-Thoms et al., 2001) cumulative dwell time, have been suggested. In this study, the threshold value of cumulative dwell time used by algorithm looking for TN locations is set up to *330ms*. A proper choice of the clustering procedure should fix with the applied model. Therefore, the unlimited number of decision sites excludes the fuzzy approach to clustering algorithms (Theodoridis et al., 1999). The spatial distribution of fixations was used to build a fixation map (Wooding, 2002), where each point is represented by the 3-D Gaussian function with sigma value related to the special accuracy of eye tracking measurements ( $\sigma = 0.5$  visual angle) (**Figure 2**).



**Figure 2.** An example of the hotspot plot created based on the fixation distribution spread through image plane collected from one subject. The most attractive regions (dwelt over for 1sec) are represented by brightness spots. The red spots indicated short gazes (around 300ms). The 3 red crosses marked the observer's overt decisions. The 5 nodule-contained ROI are shown within boxes. In this example, observer made 1 TP, 2 FP, 4 FN and several TN.

A fixations map with mouse click decision points and nodule locations are used to find sites where a particular type of decision has been made. The paradigm of the decision outcome classification originated from signal detection theory (Dorfman et al., 1996; Jiang et al., 1996), according to which the observer can make one of four possible types of decision per site:

- true-positive TP, where a nodule-containing region has been marked by observer;
- false-negative FN, where the observer missed the nodule-containing region;
- true-negative TN, is the dwelled and unmarked normal region,
- false-positive FP, where the observer marked a normal region.

The TN locations are found through the iteration peak-seeking algorithm applied into the fixation map with the 330ms threshold of the dwell time. Each iteration step results with the highest peak location. That peak coordinates is recorded as a TN location. The following step neglects all previous recorded peaks considering the only that data which has not contributed to already found locations. The algorithm carries on until the last peak over the threshold.

#### **2.4. Wavelet approach to spatial frequency analysis**

Neurophysiological and psychophysical findings show that multiscale transforms seem to take place in the visual cortex of mammals (Hubel, 1995), such that visual perception depends to some extent on the frequency components of the modified contrast (Cambell et al., 1965) and the orientation of the stimuli (Cambell et al., 1966). Hence, humans perceive the world through a number of visual channels such as colour and contrast, but some information is carried by independent spatial frequency ranges and orientations (Cambell et al., 1967). Recent studies have shown that both local and global image-based elements have particular physical properties that are correlated with the performance and the level of experience of human observers in cancer detection (Mello-Thoms, 2006). The spatial frequency analysis of particular sites gives a view inside the processes of visual perception in a radiological task of lung nodule detection from chest radiographs (Mello-Thoms et al., 2002). Although several psychophysical studies have shown the importance of signal decomposition in several frequency channels, there still is no statistical model to combine the information provided by the different channels. Multi-resolution analysis (MRA) approach to that task uses the wavelet transformation, which application in vision research has been influenced by the fact that the methodology tries to mimic the brain processing algorithms analysing visual stimuli in time order from fine, holistic to coarse, local details (Mallat, 1996). Active vision strategies compensate for the non-uniformity of visual resolution by moving the fovea with eye saccades. Regions of a scene with high information content are scanned successively. These saccades are partly guided by the lower resolution information gathered at the periphery of the retina. In consequence, the radiologists' visual system may operate on the local and global spatial frequency features and influence on image searching and interpretation, even without observer awareness of such features (Krupinski, 2003). The multiscale decomposition gives an intermediate representation between Fourier and spatial representations, where the local regularity of  $f(x)$  can be characterized based on the wavelet orthogonal basis expansion and can be interpreted as an analysis of the signal in a set of independent frequency channels in the human visual model (Marr, 1982).

Wavelet decomposition exploits the local physical properties, by which transformation is applicable for the non stationary signals analysis, i.e. for radiographs. The local background around a particular location seems be crucial in the visual searching and decision-making process (Mello-Thoms, 2006). The locality of these regions of interest (ROI) is defied by the foveal visual angle (which is approximately a 100 pixels box viewed from 60-70cm). In this way, image processing techniques were implemented to analyse the physical properties of local background selected locations – at decision-point sites. In particular, the visual information, as signal energy, carried by spatial frequencies  $\omega(\omega_x, \omega_y)$  in a certain orientations  $\theta = \tan^{-1}(\omega_x/\omega_y)$  are the main interest of this research. The Wavelet Packet Transformation (WPT) algorithm (Mallat, 1996) has been applied to spatial frequency analysis of selected ROIs up to the 3<sup>rd</sup> level of decomposition frames by Daubechies functions. This band-pass filtering iteration algorithm is designed to perform hierarchical transformations, where obtained from the previous level ( $j-1$ ) sub-images are transforming again as an input signals at level  $j$  and forward to further analysis until the last level of decomposition (**Figure 3**). Each sub-image  $f^j(x,y)$  from a particular level  $j$  represents different information called a spatial frequency band (SFB), and so, the approximations were obtained as a low spatial frequency representation of the input ROI:

$$a^j(x,y) = \langle f^{j-1}(x,y), \phi(x)\phi(y) \rangle$$

(2)

whereas high spatial frequency components are contained by sub-images represented details in particular orientations:

$$\Sigma d_i^j(x,y) = \langle f^{j-1}(x,y), \psi(x)\phi(y) \rangle + \langle f^{j-1}(x,y), \phi(x)\psi(y) \rangle + \langle f^{j-1}(x,y), \psi(x)\psi(y) \rangle$$

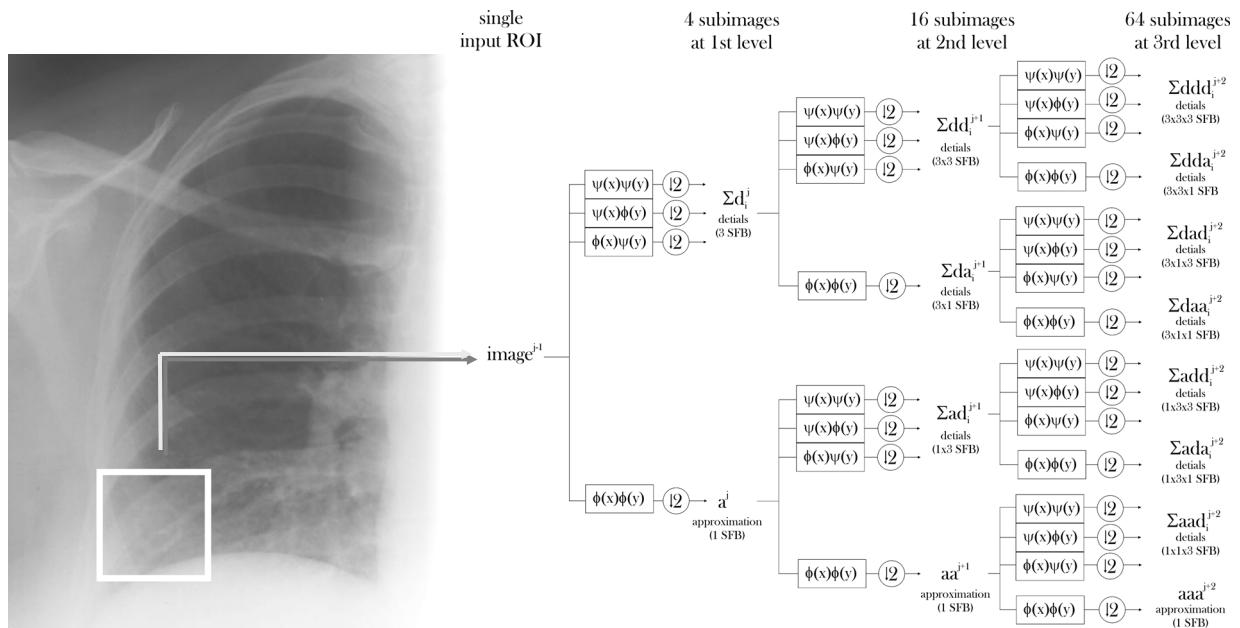
(3)

As a result, 84 bands were obtained for each ROI. Each band has a physical meaning and contains a certain amount of information, which has been quantified according to the logarithm of energy within that sub-image:

$$E = \log(\Sigma_{n,m} |f(x_n, y_m)|^2)$$

(4)

That value per band was used in further statistical analysis. The amount of texture similarities was quantified according to the ANOVA of energy within SFB considering the type of decision and the author level of experience at the statistical significance  $p < 0.05$  with Scheffe post-hoc test (Mello-Thoms et al. 2002). The number of SFB with significance difference between particular factors is interpreted as a dissimilarities indicator. The more bands are different, the less similar are two compared representations.



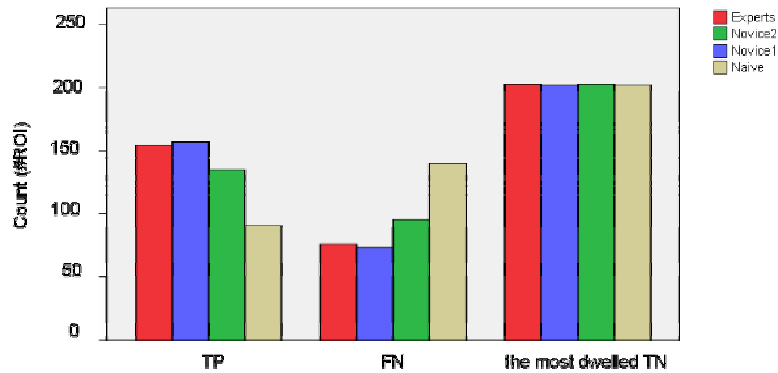
**Figure 3.** Schema of the Spatial Frequency Analysis of selected ROI. Wavelet Packet Transforms up to 3<sup>rd</sup> level of decomposition were applied to frame the input 2D signal into 4+16+64 spatial frequency bands, each characterized by different range in spatial frequency at various orientations. Approximations of the signal were obtained from 2D low-pass filtering in both orientations (*rows*( $y$ ) and *columns*( $x$ )):  $f(x)f(y)$ , whereas details obtained from high-pass  $y$  filtering:  $f(x)y(y)$ ,  $y(x)f(y)$ ,  $y(x)y(y)$ , represent high frequency components at particular scale and orientation.  $\sqrt{2}$  indicates the down-sampling step in transformation algorithm.

### 3. Results and discussion

#### 3.1. Decision sites

A post recording procedure has been applied to the eye tracking data (§2.3). First, valid samples were filtered from saccades and used to localize fixations through spatial-time clustering. Second, the fixation map concept has been used to find the locations of the decision sites. Based on the relative distances between the true nodule locations, observer mouse marks and centres of the fixation clusters over 300ms the set of decision sites have been pointed out. In this research, the selected regions of interest have been considered as follow: all TP, FN and the most dwelled TN sites from all observers and cases data (Figure 4). Our question focused on physical image differences between:

1. regions that attract visual attention and are correctly identified as containing a lesion and
2. regions that are correctly identified as containing *no lesion but attracted sustained visual attention*



**Figure 4.** The populations of selected regions (TP, FN and the most dwelled from each case TN decisions sites). The combined TP and FN sites represent the nodule-contained regions perceived by each observer. The most dwelled TN sites have been selected according to the cumulative dwell time value bigger than 99.22% of the range frames by the less and the most dwelled TN in a particular case.

### 3.2. Spatial frequency features

Locations have been chosen and analysed using 2D Wavelet Packet Transform up to the 3<sup>rd</sup> scale (level of decomposition) framed by Daubechies wavelet functions in order to answer for the research questions (**§1**). The zero-padding extending procedure has been implemented in the decomposition algorithm which extended the image boundary (Misiti et al., 2008). The statistical comparison between the spatial frequency features was performing according to the General Linear Model. An analysis of variances between the mean values of energy carries by spatial frequency bands was used to find out the statistical significance at  $P < 0.05$  levels between the different decision locations.

The population of feature vectors related to the TP sites has been compared with the set of vectors related to the most dwelled TN site from the case where observer committed at least one TP response. The number of spatial frequency bands across 3 scales of decomposition and 4 levels of observer experience, showing those with significance difference is summarised in the **Table 3**.

**Table 3.** Summaries of the results of the statistical analysis of variance spatial frequency bands up to the 3<sup>rd</sup> scale which demonstrated a significance difference at  $P < 0.05$ . The totals show how many bands carry significance difference between TP and TN zones in terms of number of bands per scale. The percentages presented in the last column show the proportion of pointed bands overall representations.

Observer Groups of Experience Level	Scale 1	Scale 2	Scale 3	total	
<b>Naive</b>	1	6	36	<b>43</b>	51.19%
<b>Novice1</b>	1	9	42	<b>52</b>	61.90%
<b>Novice2</b>	2	12	51	<b>65</b>	77.38%
<b>Experts</b>	1	9	49	<b>59</b>	70.24%

The 3 scale analysis is an attempt to find out which spatial frequency properties have been tuned, in order to correctly distinguish the suspicious object as a truly normal or abnormal region in terms of lung nodule recognition. The trend in the dynamic of the statistical nature of particular bands has been observed across radiological experience. However it is still too early to report any conclusive correlation or observed trend with the level of expertise.

### 3.2. Correctness response as a function of previous committed overt decision

The probability of an error event as a function of the correctness of a previous outcome is studied based on the number of overt decisions made in a certain time order. The decision order was measured by the order of mouse clicks events registered during eye tracking experiments where subjects used it for marking suspected ROIs. The number of TP and FP responses followed by an incorrectly marked region is presented at upper plot (**Figure 6**). Because, the chance that a subject will make another FP after a FP is much bigger than a correct decision after a FP, the probability of mistakes increase as a function of the correctness of the previous response as follows:

$$p(FP|FP) > p(TP|FP)$$

(5)

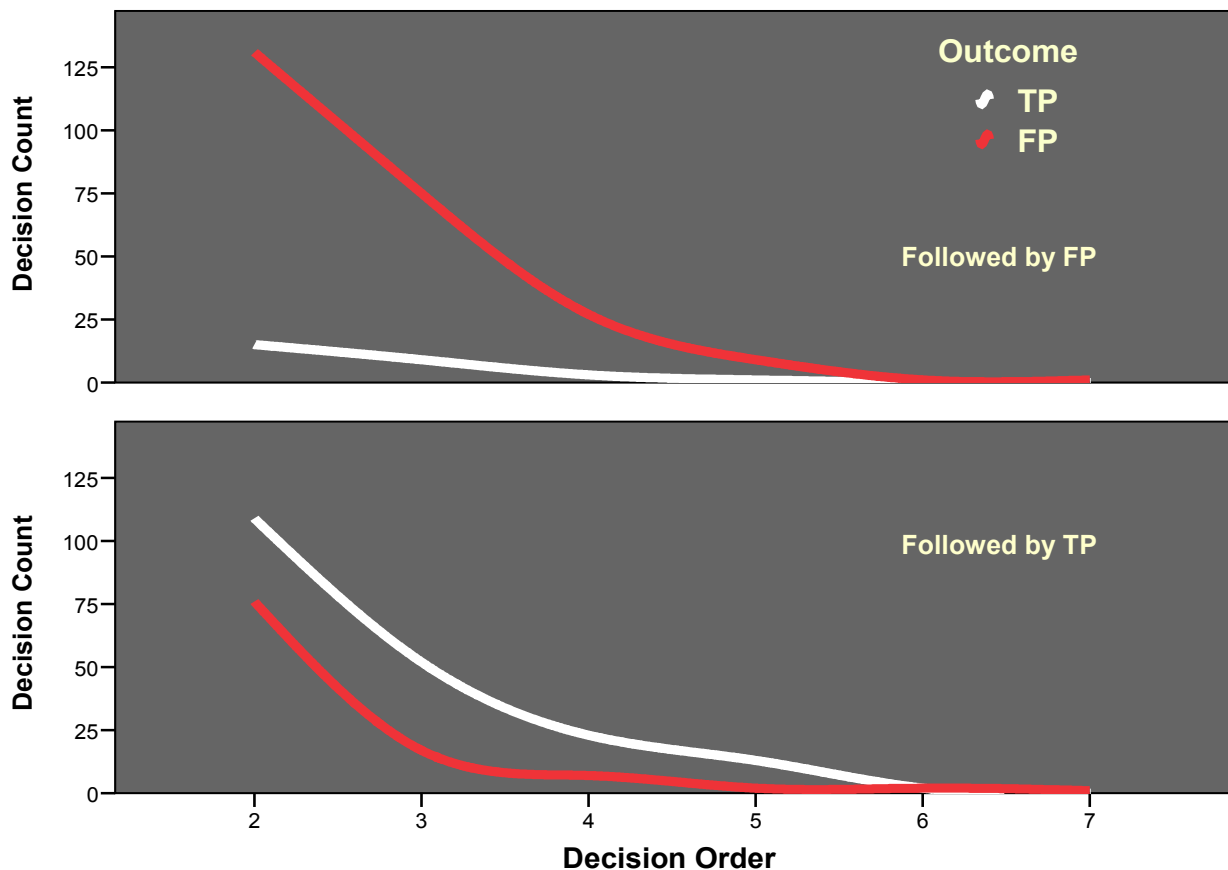
Similarly, a correctly detected and localized pulmonary nodule is more likely to be identified if it follows a correctly recognised and detected other nodule-contain region:

$$p(TP|TP) > p(FP|TP)$$

(6)

These results are consistent with recent published perceptual studies in a different imaging modality (Mello-Thoms, 2009).





**Figure 6.** Overt decisions (TP and FP) distribution across time order. Upper plot shows the number of overt decisions (y axis) followed by False Positive, whereas lower plot depicts the number of TP and FP decisions followed by True Positive.

#### 4. Conclusions

This study is an extension of the perceptual approach to lung nodule detection studies in chest imaging (Manning et al., 2004; Krupinski, 2004) by applying spatial frequency analysis, and is also comparable with a similar study from mammography (Mello-Thoms et al., 2002). The correlations between the physical properties of the sites where a particular decision has been made, and the type of the response, have been analyzed. We have shown that there are spatial frequency and orientation characteristics in image regions that contain lesions that are replicated in non-lesion areas that attract visual attention. Experts perform better in terms of the nodule detection, but also, they allocate more visual attention on those regions which are similar to the nodules in terms of spatial frequency. The expert seems to perform better because of their ability to distinguish between TN and TP despite any spatial frequency similarities. Therefore, the spatial frequency analysis gained insight into that phenomenon, showing the set of features based on those two types of regions which have been perceived differently. The results obtained so far show the correlations between the spatial frequency properties of the local regions and the type of decisions made on it, and also the spatial distribution of visual attention. The correlations between levels of experience in the radiological task are a work still in progress. This interpretation shows some agreement with the conclusions presented by Mello-Thoms et al., who suggest that experts may use specific neural connections – a spatial frequency channels tuned to specific object detection – during visual searching in a radiological task.

Furthermore, the importance of the holistic view on the radiological performance is studied through the analysis of the probability of an error event as a function of the previous outcome correctness is studied. A TP is more likely to happen if it is preceded by a correctly recognised and detected nodule-contain region than if it follows an incorrectly marked normal ROI. By analogy, an FP occurs more frequently after FP than after TP. According to the internal search template view (Desimore et al., 1995), the subject's visual attention is guided by the mental representation (an internal template) of the target (pulmonary nodule). Moreover, the internal

template is specific to the visual learned target's features (Bravo et al., 2009). Hence, when the template is set up upon a first overt decision, it is more likely that the following decision will be similar in terms of perceptual features. Therefore, if the first decision was correct, that TP template is matched more efficiently with another nodule-contain ROI, which is demonstrated by the visual attention allocation in the similar type of the following outcome. Similar situations may occur in FP cases. Our work is confirmation of those similar effects which have been found previously in a different modality (Mello-Thoms, 2004, 2009). Even though radiological skills are not necessarily transferable, the principals that rule the perceptual-cognitive interaction seem to be very similar.

## References

- BRAVO, M. & FARID, H. (2009) The specificity of the search template. *Journal of Vision*, **9(1):34**, pp. 1-9.
- BRENNAN, P.C., MCENTEE, M., EVANOFF, M., PHILLIPS, P., O'CONNOR, W.T. & MANNING, D.J. (2007) Ambient Lighting: Effect of Illumination on Soft-Copy Viewing of Radiographs of the Wrist. *American Journal of Roentgenology*, 188, p. W177.
- CAMBELL, F.W. & GREEN, D. (1965) Optical and retina factors affecting visual resolution. *Journal of Physiology (London)*, 181, pp. 576-593.
- CAMBELL, F.W. & KULIKOWSKI, J. (1966) Orientation selectivity if the human visual system. *Journal of Physiology (London)*, 197, pp. 437-441.
- CAMBELL, F.W. & ROBSON, J.G. (1967) Application of Fourier analysis to the visibility of gratings. *Journal of Physiology (London)*, 197, pp. 551-566.
- CARMODY, D.P., NODINE, C.F. & KUNDEL, H.L. (1981) Finding lung nodules with and without comparative visual scanning. *Perception and psychophysics*, 29, p. 594.
- CHAKRABORTY, D.P. (2006) A search model and figure of merit for observer data acquired according to the free-response paradigm. *Physics in Medicine and Biology*, 51, pp. 3449-3462.
- DESIMORE, R. & DUNCAN, J. (1995) Neural mechanism of selective visual attention. *Annual review of neuroscience*, 18, pp. 193-222.
- DOFMAN, D., BERBAUM, K.S., METZ, C.E., LENTH, R.V., HANLEY, J.A. & DAGGA, H.A. (1996) Proper Receiver Operating Characteristic Analysis: The Bigamma Model. *Academic Radiology*, 4, pp. 138-149.
- DONOVAN, T. (2008) Performance changes in wrist fracture detection and lung nodule perception following the perceptual feedback of eye movements. *School of Medical Image Sciences*, Lancaster, Lancaster University.
- ETHELL, S.B. (2003) Disease prevalence: an influential factor in radiological performance. *School of Medical Imaging Sciences*, Lancaster, Lancaster University.
- GUR, D., ROCKETTE, H.E., ARMPFIELD, D.R., BLACHAR, A., BOGAN, J.K., BRANCATELL, G., BRITTON, C.A., BROWN, M.L., DAVIS, P.L., FERRIS, J.V., FUHRMAN, C.R., GOLLA, S.K., KATYAL, S., LACOMIS, J.M., MCCOOK, B.M., THAETE, F.L. & WARFEL, T.E. (2003) Prevalence effect in a laboratory environment. *Radiology*, 228, p. 10.
- HOFFMAN, J.E. (1998). Visual attention and eye movemnets. In PASHIER (ED.), *Attention*. Hove, England, Cambrige University Press.
- HUBEL, D. (1995) *Eye, Brain, and Vision*. , <http://hubel.med.harvard.edu>.
- IRWIN, D.E. (1992). Visual Memory Within and Across Fixations. In K. RAYNER (ED.), *Eye movements and visual cognition: Scene perception and reading*. New York, Springer-Verlag.
- JIANG, Y., MENZ, C. & NISHIKAWA, R. (1996) A Receiver Operating Characteristic Partial Area Index for Highly Sensitive Diagnostic Tests. *Radiology*, 201, pp. 745-750.
- KRUPINSKI, E. (2004) Pulmonary nodule detection: what features attract attention?. *SPIE Medical Imaging: Image Perception, Observer Performance, and Technology Assessment*, 5372, pp. 122-127. , SPIE Library.

- KUNDEL, H.L. & NODINE, C.F. (2004) Modelling visual search during mammogram viewing. Chakraborty, D.P. & Eckstain, M.P. (Eds.) *SPIE Medical Imaging: Image Perception, Observer Performance, and Technology*, 5372, pp. 110-115. San Diego, CA, SPIE - The International Society for Optical Engineering.
- KUNDEL, H.L., NODINE, C.F. & KRUPINSKI, E.A. (1990) Computer-displayed eye position as a visual aid to pulmonary nodule interpretation. *Investigative Radiology*, 25, p. 890.
- KRUPINSKI, E.A. (2003) The future of image perception in radiology: synergy between humans and computers. *Academic Radiology*, 10, p. 1.
- MALLAT, S. (1996) Wavelets for a vision. *IEEE*, 84, pp. 604-614. , .
- MANNING, D., GALE, A. & KRUPINSKI, E. (2005) Perception research in medical imaging. *The British Journal of Radiology*, 78, pp. 683-685.
- MANNING, D.J., ETHELL, S.C. & DONOVAN, T. (2004) Categories of observer error from eye-tracking and AFROC data. Chakraborty, D.P. & Eckstain Miguel P (Eds.) *SPIE Medical Imaging: Image Perception, Observer Performance, and Technology*, 5372, pp. 90-99. San Diego, CA, SPIE Library.
- MANNING, D.J., ETHELL, S.C., DONOVAN, T. & CRAWFORD, T. (2006) How do radiologists do it? The influence of experience and training on searching for chest nodules. *Radiography*, 12, p. 134.
- MARR, D. (1982) *Vision*. New York, Freeman.
- MELLO-THOMS, C. (2006) The problem of image interpretation in mammography: effects of lesion conspicuity on the visual search strategy of radiologists. *The British Journal of Radiology*, 000, pp. 1-6.
- MELLO-THOMS, C. (2006) How does the perception of a lesion influence visual search strategy in mammogram reading?. *Academic Radiology*, 13, pp. 275-288.
- MELLO-THOMS, C. (2009) The 'Holistic Grail': Possible implications of an initial mistake in the reading digital mammograms. Sahiner, B. & Manning, D.J. (Eds.) *SPIE Medical Imaging: Image Perception, Observer Performance, and Technology Assessment*, 7263-20, p. 20. Orlando, FL, SPIE Library.
- MELLO-THOMS, C. & CHAPMAN, B. (2004) A preliminary report on the role of spatial frequency analysis in the perception of breast cancers missed at mammography screening. *Academic Radiology*, 11, pp. 894-906.
- MELLO-THOMS, C., DUNN, S.M., NODINE, C.F. & KUNDEL, H.L. (2001) An analysis of perceptual errors in reading mammograms using quasi-local spatial frequency spectra. *Journal of Digital Imaging*, 14, pp. 117-123.
- MELLO-THOMS, C., DUNN, S.M., NODINE, C.F., KUNDEL, H.L. & WEUNSTEIN, S.P. (2002) The perception of breast cancer: what differentiates missed from reported cancers in mammography?. *Academic Radiology*, 9, pp. 1004-1012.
- MISITI, M., MISITI, Y., OPPENHEIM, G. & POGGI, J. (2008) *Wavelet Toolbox™ User's Guide*. www.mathworks.com, The MathWorks, Inc.
- NODINE, C.F. & KUNDEL, H.L. (1987) Using eye movements to study visual search and to improve tumour detection. *Radiographics : a review publication of the Radiological Society of North America, Inc*, 7, p. 1241.
- ROBINSON, D.A. (1968) The oculomotor control system: a review. *Proceedings of the IEEE*, 56(6), pp. 1032-1049. , IEEE Press.
- SHEBILSKA, W.L. & FISHER, D.F. (1983). Understanding extended discourse through the eyes: how and why. In R. GRONER, C. MENTZ, D.F. FISHER & R.A. MONTY (EDS.), *Eye Movements and Psychological Functions: International Views*. Hillsdale, NJ, Lawrence Erlbaum.
- THEODORIDIS, S. & KOUTROUMBAS, K. (1999) *Pattern Recognition*. San Diego, CA, Elsevier Science Academic Press.
- WOODING, D.S. (2002) Eye movements of large populations: II. Deriving regions of interest, coverage, and similarity using fixation maps. *Behavior Research Methods, Instruments, & Computers*, 34(4), pp. 518-528.

## A Clockwork Orange: A Unique Challenge for the Translator

**Ines Radionovas**, *School of Languages, University of Salford*

The profession of the translator being often underestimated, one major interest in translation studies should be to illustrate the complexity of the translation process. Amongst the most intriguing examples of extraordinary challenges for the translator are certainly literary texts that play with language. Often, this play is not merely an artistic exercise but rather an essential part of and support for the story itself. It can be said that as soon as the form supports the content, translation reaches its limits and the translator has to choose which aspect to live with the loss of. One example for this phenomenon is Anthony Burgess' novel *A Clockwork Orange*.

The novel was first published in 1962. It tells the story of Alex and his friends, juvenile delinquents who shoplift, mug people and take drugs. In the first part of the novel, the life of the gang is described. The friends, for instance, break into the house of a writer, beat him up and rape his wife. One day, however, the others turn against Alex and leave him to get caught by the police; Alex is sentenced to 14 years in prison. The second part is set in prison, where Alex is offered the opportunity to take part in an experiment that would possibly allow him to be released early. This experiment can be compared to Pavlov's conditioning; Alex has to watch violent videos while being given drugs that make him feel nauseous. As a result, he feels sick as soon as he even thinks about violence. He is released from prison.

In the third part, Alex is abandoned by his parents and beaten up in the street, unable to defend himself. He seeks refuge in the house of the writer he has once broken into. This writer is revealed as having taken part in a campaign against the conditioning treatment Alex had undergone. When the writer recognises him, he drives Alex to commit suicide in order to show the negative results of the treatment. Alex survives and is cured of the conditioning effects. In the end, he decides to change his life.

The linguistic particularity of *A Clockwork Orange* is the language that Alex and his friends speak. It is an invented youth slang called *Nadsat*, which involves mainly words derived from Slavic languages, but also rhyming slang, onomatopoeia and Shakespearean language. The following short excerpt is to illustrate *Nadsat*:

*WHEN we got outside of the Duke of New York we **viddied** by the main bar's long lighted window, a burbling old **pyahnitsa** or drunkie, howling away at the filthy songs of his fathers and going **blerp blerp** in between as though it might be a filthy old orchestra in his stinking rotten guts. One **veshch** I could never stand was that. I could never stand to see a **moodge** all filthy and rolling and burping and drunk, whatever his age might be, but more especially when he was real **starry** like this one was. He was sort of flattened to the wall and his **platties** were a disgrace, all creased and untidy and covered in **cal** and mud and filth and stuff. So we got hold of him and cracked him with a few good **horrorshow tolchocks**, but he still went on singing. (...)* (Burgess, 1962: 14)

Glossaries have been compiled that show the meanings and origins of *Nadsat* terms. Here a short excerpt:<sup>1</sup>

<b>Nadsat</b>	<b>Meaning</b>	<b>Origins</b>
<i>appy polly loggy</i>	apology	school boy speak
<i>baboochka</i>	old woman	Russian: babooshka (grandmother)

<sup>1</sup> Compiled from Hyman (1963) and Katy Albrecht's glossary in Burgess (1997), adapted.

<i>xbaddiwad</i>	bad	school boy speak
<i>banda</i>	band	Russian: banda (band, gang)
<i>bezoomny</i>	mad	Russian: byezoomiyi (mad, insane)
<i>Biblio</i>	library	Russian: biblioteka (library)
<i>bitva</i>	battle	Russian: bitva (battle)
<i>Bog</i>	God	Russian: Bog (God)
<i>bolnoy</i>	sick	Russian: bolnoy (sick)
<i>bolshy</i>	big	Russian: bolshoy (big)
<i>bootick</i>	shop	French: boutique
...	...	...

The function of this artificial slang is threefold. Anthony Burgess explains that it is meant to brainwash the reader, as he or she is forced to learn this language while reading, just like Alex is brainwashed in the story. Secondly, *Nadsat* is a means of describing violence and letting the characters curse, without being too explicit, thus keeping the story aesthetic. The third function of *Nadsat* is to hide the time and the place of the story. Had Burgess chosen any kind of existing slang, the story would be fixed to the period and region where this slang was spoken. With *Nadsat*, the story could have happened anytime, anywhere, thus making the moral message of the novel eternally valid.

This moral message is that, in Burgess' opinion, it is always better for humans to have the freedom of mind to decide for themselves whether they want to be good or bad. Creating a peaceful society by taking this freedom away can never be an acceptable solution. This point of view that the author wants to convey makes the novel and its form so important. Moreover, the story also has an autobiographical character, as Burgess has included himself in the novel as the writer, who is robbed and violated in his own house. Burgess and his wife were really once victims of a violent robbery and he blamed this incident for her early death. This background illustrates how carefully any translator has to treat this text in order to be faithful to the author.

The following translation analysis exemplifies translation strategies used in three different versions in two languages in conveying an invented language. *A Clockwork Orange* was first translated into German in 1972 by Walter Brumm. A second translation, by Wolfgang Krege, was published in 1997 and is the only version that is available on the market today. This retranslation obviously raises the question as to why the publishing house decided to publish a new version and what are the differences compared to the old translation. In any case, the retranslation was badly received and Krege, the infamous retranslator of *The Lord of the Rings*, received harsh criticism. The third translation that is analysed is the French version by Georges Belmont and Hortense Chabrier from 1972.

The task for the translator of this novel is to transfer *Nadsat* into the target language without creating any unwanted associations in the target reader or making the whole translation sound too strange and therefore strenuous to read. Furthermore, *Nadsat* does not only consist of separate neologisms. These neologisms are often integrated into English idiomatic expressions, thus presenting the translator with an additional linguistic problem. As methodology for this analysis, different strategic models are applied. Vinay & Darbelnet's seven "procedures" of translation are also part of Chesterman & Wagner's more detailed list of "strategies", as they call them. For Jean-Pierre Mailhac, the difference between the terms is that a procedure is a "means of translating a particular element as part of a strategy", whereas a strategy is "conceptually broader". One example of a strategy could be the decision to produce a domesticated or a foreignised translation.

Usually, the decision to be made is either to produce an exotic text (foreignised) that still has the flavour of the source language, or to make the translation sound like an original text (domesticated). In the case of *A Clockwork Orange*, it is relatively obvious that the exoticism that the source text already has needs to be maintained. However, both German translators do not consistently stick to this strategy. Another question that arises is whether the aspects of spelling and idiomatic expressions, both largely influenced by the English language in the original, should be kept as they are or adapted to the target language. Selected examples of procedures on the word-level will give insight into the work of the three different translators.

The first procedure is probably the most obvious means of maintaining the foreignness of the original. Wherever possible, **borrowing** seems to be the best decision in the case of Nadsat. However, sometimes slight alterations might be necessary. First some examples of **exact borrowing**, where it occurs in all three translations (the numbers in the first column indicate the part and the chapter):

	<b>English</b>	<b>German</b>	<b>French</b>	<b>Meaning</b>
I.1	<i>litso</i>	<i>Litso</i>	<i>litso</i>	face
I.2	<i>maslo</i>	<i>Maslo</i>	<i>maslo</i>	butter
I.3	<i>sabog</i>	<i>Sabog</i>	<i>sabog</i>	shoe

It is not a coincidence that this list only contains nouns, as nouns seem to be the easiest to borrow, given that they require hardly any morphological changes.

There are also cases where both German translators borrow the same term but with slight **changes in spelling** (underlined):

	<b>English</b>	<b>German</b>	<b>French</b>	<b>Meaning</b>
I.2	<i>devotchka</i>	<i>Dew<u>o</u>tschka</i>	<i>d<u>e</u>votchka</i>	girl
I.1	<i>goloss</i>	<i>Golo<u>ß</u></i>	<i>golo<u>s</u>se</i>	voice
II.2	<i>keeshkas</i>	<i>K<u>i</u>schkas</i>	<i>k<u>i</u>shkas</i>	guts
II.1	<i>prestoopniks</i>	<i>Prest<u>u</u>pniks</i>	<i>prest<u>o</u>upnick</i>	criminals

These examples illustrate where spelling changes are necessary. In order to maintain the original pronunciation of the neologism, the French sometimes requires an accent on the "e" or a mute "e" at the end. The "ee" has to be turned into an "i" in both target languages, whereas the "oo" has to become a "u" in German and an "ou" in French. These can be considered as phonological domestication. Spelling domestication can be observed in the first two German examples, where the change from "v" to "w" and from "ss" to "ß" is not necessary for the pronunciation.

Of course, the two German translators have not always taken the same decisions. The following table shows examples where both translators borrow terms, but with different spellings:

	<b>English</b>	<b>Brumm</b>	<b>Krege</b>	<b>Meaning</b>
I.1	<i>slovo</i>	<i>Slovo</i>	<i>Sl<u>o</u>vo</i>	word
I.3	<i>noz</i>	<i>Noz<u>h</u></i>	<i>Nosch</i>	knife
II.1	<i>yahoodies</i>	<i>Yahoodies</i>	<i>J<u>a</u>h<u>u</u>dis</i>	Jews

These examples can already be used to draw comparisons between the two German versions. Walter Brumm seems to stick more to the original, whereas Wolfgang Krege domesticates the

spelling more. Thereby he makes the pronunciation clearer – even clearer than in the original, as the pronunciation of “*noz*h” is all but obvious to the English reader. Walter Brumm, on the other hand, can be reproached with inconsistency, as in different words, he takes different spelling decisions for the same letters/sounds: he spells “*Slovo*” as in the original, but “*Dewotschka*” with “w” and he keeps the “oo” in “*Yahoodies*”, but domesticates the spelling in “*Prestupniks*”.

All in all, both German translators borrow *Nadsat* terms, mostly nouns, but not the majority of them. Although Walter Brumm seems to be more inconsistent in his spelling adaptation, the same problems can be found in Wolfgang Krege’s version. The French translators, on the other hand, borrow the vast majority of *Nadsat* terms. Some of them are adopted as they are, others with slight spelling changes. These domesticating changes in spelling, however, are consistent throughout all the borrowed words.

The second lexical procedure seems, besides borrowing, the other possibility of remaining faithful to the *Nadsat* vocabulary used in the original. **Lexical creation** means that the translator invents a neologism of their own to replace a term in the original. This may well be considered just as good as copying the original term, in terms of faithfulness. Within this procedure, it can be observed that the translator often chooses neologisms that raise certain associations, as can be seen in the first two examples in the following table (the terms in bracket are not lexical creations and only stated for the sake of completeness):

	English	Brumm	Krege	French	Meaning
I.4	<i>bitva</i>	<i>Rampferei</i>	<i>Geteuse</i>	[ <i>bitva</i> ]	battle
I.6	<i>kots</i>	<i>Murken</i>	<i>Kotter</i>	<i>mitons</i>	tomcats
II.1	<i>horrible pishcha</i>	<i>Schlengen</i> -	<i>Schlengen- fraßes</i>	[ <i>l'horrible Pischa</i> ]	food

Possible associations are “*Rauferei*” (fight) for “*Rampferei*”, “*Getöse*” (noise) for “*Geteuse*” and “*Kater*” (tomcat) for “*Kotter*”. A very interesting case is the neologism “*mitons*” in French, probably a combination of “*minou*” and “*chaton*”, both meaning “cat”. With these derivations, the target term becomes more understandable than the source term; however, this compensates for cases where the reverse applies.

In the French version, borrowing is definitely the leading procedure, therefore lexical creation is not as frequent as in the German versions. However, here are some more French examples:

	English	French	Meaning
III.7	<i>cal</i>	<i>gouspin</i>	shit
I.1	<i>cancers</i>	<i>cancerettes</i>	cigarette
I.4 I.5	<i>cutter</i>	<i>lollypop mouizca</i>	money

It should be mentioned here that, unlike the French translators, both German translators fail to provide an adequate solution for “*cal*”. Given one of the aforementioned functions of *Nadsat*, namely hiding strong language, it can be considered a major faux-pas in terms of faithfulness to the original that the word “*Scheiße*” can be found frequently in the two German versions, whereas Burgess does not use “*shit*” a single time.

This leads to the next procedure to be mentioned, namely **literal translation**. At first sight, literal translation is not a good choice for *Nadsat* terms, as with every term translated literally the strangeness of the original is compromised. However, there is one possibility of avoiding complete loss, which is literal translation with a **slang** expression. Burgess uses slang himself, so this form of literal translation can be considered acceptable. Nevertheless, it is only the German translators who use it, not the French.

	<b>English</b>	<b>Brumm</b>	<b>Krege</b>	<b>Meaning</b>
I.1	<i>peet</i>	<i>pitschen</i>	<i>pitschen schlürfen</i>	to drink
I.1 I.5	<i>pretty/pr etty polly</i>	<i>Moos Pulver</i>	<i>Heu Kies</i>	money
I.1 II.4	<i>rookers</i>	<i>Flossen Krallen</i>	<i>Rucken</i>	hands
I.1 I.2	<i>rot</i>	<i>Schnauz e</i>	<i>Flappe Rachen</i>	mouth

Sometimes, it is hard to distinguish between slang and neologism, as “*pitschen*” and “*Pulver*”, for instance, seem to be common in certain regions and unknown in others.

However, unfortunately the example of “*cal*” translated literally into everyday words is not the only one in the German versions. In the process of preparing this analysis, Burgess’ *Nadsat* has been found to consist of an estimated 243 terms – be it neologisms, onomatopoeia or otherwise unusual expressions. In Walter Brumm’s translation only 116 could be found and in Wolfgang Krege’s version as few as 108. Looking at these numbers, it seems as though the degree of loss is over 50% in German. The concept of loss in translation studies, however, goes hand in hand with the concept of compensation. Obviously, the decreased vocabulary volume in German cannot be dismissed, but there are instances of compensation that make up for this to a certain extent.

The following tables show two of the main kinds of **compensation**. The first kind of compensation is the **replacement** of a straightforward term in the original with a *Nadsat* term. Even Anthony Burgess is inconsistent in his novel and sometimes uses the normal English expression instead of the existing *Nadsat* term. The three examples from Brumm’s version show instances where the translator uses these inconsistencies in the original to compensate for loss somewhere else in the text:

	<b>English</b>	<b>Brumm</b>
I.3	<i>guts</i>	<i>Kischkas</i>
I.4	<i>ate</i>	<i>menkelte</i>
I.1	<i>old veck</i>	<i><u>starrigen</u> Fecken</i>

Of course, all three translators use this method, as well as the second one, which could be called **insertion**. The next table shows examples where Wolfgang Krege inserts *Nadsat* terms into the sentence, where there is none in English:



	<b>English</b>	<b>Krege</b>
I.1	<i>The four of us were dressed in the height of fashion</i>	<i>Alle vier trugen wir <u>Plattis</u>, die voll im Trend lagen</i>
I.1	<i>you're nothing but a filthy-minded old skitebird</i>	<i>Na, du bist mir ein Wüstling. Nichts als Sauereien im <u>Gulliver</u>.</i>
II.3	<i>I lay back in my bed</i>	<i>ließ ich meinen <u>Gulliver</u> ins weiche <u>Poduschka</u> zurücksinken</i>

To explain this compensating procedure in more detail: in the first example, the word "dressed" is translated by "Plattis tragen", i.e. "to wear clothes", "Plattis" being a Nadsat term. The other two examples follow the same principle.

The French translators use the same two main methods of compensation. The first example is a replacement, the second one an insertion:

	<b>English</b>	<b>French</b>
I.4	<i>mum</i>	<i>môman</i>
I.3	<i>going:</i>	<i><u>bavochant des trucs</u> genre</i>
I.1	<i>and all that chepooka</i>	<i>et autre blabla à la <u>conskov</u></i>

The last example in the table, in particular, shows their inventiveness in this respect. Derived from "à la con" (slang for something like "stupid"), they create "à la conskov", thereby even transferring the Slavic character of Burgess' Nadsat. However, as the French translators borrow most of the Nadsat terms, compensation is consequently not as frequent – not as frequently needed – in their version as in the German ones.

The last aspect of Anthony Burgess' language in *A Clockwork Orange* to be mentioned in terms of difficulties for the translator is the fact that the author even replaces words in English idioms by Nadsat terms. This is where it becomes clear that Nadsat is not only a vocabulary list, but a collection of terms that are woven into the source language. The question this raises is the degree to which this combination should be maintained in the translation, or otherwise, if the term should rather be separated from the idiomatic context and integrated in a target language idiom.

The following two examples of idiomatic expressions with integrated Nadsat terms show different solutions to this question:

	<b>English</b>	<b>Brumm</b>	<b>Krege</b>	<b>French</b>
I.2	<i>near laughed his <u>gulliver</u> off</i>	<i>lachte sich darüber fast seinen <u>Gulliver</u> ab</i>	<i><u>schmitzte</u> sich einen Ast</i>	<i>s'est presque pété le <u>gulliver</u> de rire</i>

I.4	me having a <u>sladky</u> tooth	weil ich ihn gern <u>sladki</u> habe	denn ich bin ein Süßmaul	vu que j'ai une dent <u>sladky</u>
-----	--	---	--------------------------------	---

Unsurprisingly, all three translators use various ways of dealing with these idioms. In the first example, Walter Brumm leaves "gulliver" in its context. The problem here is that in German, it is uncommon to "laugh one's head off". "Sich den Arsch ablachen" would be the common equivalent in German, which might lead to the German reader assuming that "Gulliver" stands for "bottom" (or a more vulgar expression) instead of "head". Given that Burgess wants the reader to learn *Nadsat* while reading, it cannot be in his interest that the reader is confused by misleading usage of the terms by the translator. Krege compensates for the loss of "gulliver" by replacing "laughed" with his own neologism "schmitzte". The French translators were lucky enough to find a French idiom with the word "head" in it.

This example and the second one have in common that the expression is so familiar to the English reader that it is easy for him to derive the meaning of the inserted *Nadsat* term. In Brumm's translation of the second example, this effect is lost. He keeps the *Nadsat* term, but loses the idiom. The German "weil ich ihn gern sladky habe" ("because I like it sladky") is not as easy to understand as "me having a sladky tooth", as the idiom makes it clear. However, Krege on the other hand loses the idiom as well as the *Nadsat* term, which is definitely worse. The French solution shows an example of foreignisation on the idiomatic level. Belmont and Chabrier translated the whole expression literally into French. Usually this would be considered unacceptable, however, when it comes to a text that sounds foreign even to the source reader, methods like this are probably "allowed".

This leads back to the original thesis in this paper that the great importance of the form of this novel makes it an unusual text to translate. The form is basically "strange" – a strange language with influences from foreign languages and unusual language variations. In order to remain faithful to the original, the translator has to use a "strange" language, too. It is not always possible to imitate the play on words in the original, and therefore, there is loss at some points that needs to be compensated for at others. Furthermore, it seems acceptable to sometimes break translation rules that are intended to lead to a natural text. All in all, due to the constraint of the form, the translator paradoxically enough has to become more creative and recreate the text rather than merely transfer it to the target language.

In the case of *A Clockwork Orange*, the translator has to first understand the author's intentions, i.e. the function of the form and the message it conveys. Afterwards, the translator must decide about the degree of domestication he wants to apply. Three points have been observed here that are relevant for domestication: *Nadsat* terms themselves, the spelling and therefore phonetics of words and idiomatic expressions that contain *Nadsat* terms. That most of the *Nadsat* terms should be kept rather than domesticated seems obvious, but still it is only the French translators who clearly do this. When it comes to the spelling of terms, there are phonetically necessary changes and changes that are merely "cosmetic", i.e. following spelling norms of the target language rather than pronunciation rules. This decision is most probably a matter of taste, but once the translator has decided to domesticate the spelling, he should maintain this strategy. Unfortunately, both German translators are inconsistent in this respect, contrary to the French translators. The same applies to idiomatic expressions, although consistency is not essential, as some borrowed idioms amongst others that are adapted will not confuse the reader.

As this fine line of translation decisions will inevitably lead to a number of instances of loss that can be compensated for elsewhere in the text, it is extremely difficult to measure the degree of faithfulness in terms of percentage of loss and compensation. The only means of evaluating the work of the translators would be comparison. The work for this study began with the question of why the novel was translated into German twice. After an in-depth analysis and comparison of the two German translations, another question arose. This question was if there were not another, completely different approach to translating *A Clockwork Orange*. And indeed, the French version does seem to present such an alternative approach.

The question of why the novel was retranslated into German does not seem to have any obvious answer. Both translators are inconsistent in their strategies and seem to borrow unnecessarily few *Nadsat* terms. However, the question why the second translation has received such bad criticism can be answered by the assumption that the fans of the story were used to the old version and did not like to accept linguistic changes. This theory is supported by the fact that the German movie adaptation also uses the language of the first translation. Perhaps most strikingly, Wolfgang Krege decided to change the name of one of the protagonists. Maybe this mistake alone has cost him sympathy. Other than that, the second translation is not better than the first one in any linguistically measurable way.

The French version is much closer to the original. Most of the *Nadsat* terms are borrowed and consistently domesticated in terms of spelling. It does not read as being more or less 'foreign' than the original. The language is equally aesthetic to the language of the original and cursing and violence are adequately hidden with *Nadsat*. It might be less creative, but it is certainly more faithful to the original. And it does contain some very creative ideas too, as pointed out in this analysis. All in all, it can be said that the German translators have tried to act as creative writers themselves, whereas the French translators produced a very close rendition of the original. Both approaches are acceptable and it is very much a question of taste which one is to be preferred. Therefore, literary translation must be considered a form of art, where many things are possible and little is forbidden. However, whether Anthony Burgess himself would be happy with the German versions remains obscure.

## References

### Primary

- Burgess, A., (1962). *A Clockwork Orange*. Harmondsworth: Penguin.  
 Burgess, A., (1972). *L'orange mécanique*. Translated by G. Belmont and H. Chabrier. Paris: Éditions Robert Laffont.  
 Burgess, A., (1972). *Uhrwerk Orange*. Translated by W. Brumm. Munich: Heyne.  
 Burgess, A., (1997). *Clockwork Orange*. Translated by W. Krege. Munich: Heyne.

### Secondary

- Baker, M., (1992). *In Other Words: A Coursebook on Translation*. London & New York: Routledge.  
 Burgess, M., (2001), *Routledge Encyclopaedia of Translation Studies*. London & New York: Routledge.  
 Bassnett, S. & Lefevre, A., (1998). *Constructing Cultures – Essays on Literary Translation*. Clevedon: Cromwell Press.  
 Boase-Beier, J. & Holman, M. ed., (1999). *The Practices of Literary Translation - Constraints and Creativity*. Manchester: St. Jerome Publishing.  
 Boase-Beier, J., (2006). *Stylistic Approaches to Translation*. Manchester: St. Jerome Publishing.  
 Chesterman, A. & Wagner, E., (2002). *Can Theory Help Translators?* Manchester: St. Jerome Publishing.  
 Dix C. M., (1971). *Writers and their work – Anthony Burgess*. Harlow: Longman.  
 Hatim, B. & Mason, I., (1990). *Discourse and the Translator*. London: Longman.  
 Hatim, B. & Munday, J., (2004). *Translation - An advanced resource book*. London & New York: Routledge.  
 Holmes, J. S., (1994). *Translated!- Papers on Literary Translation and Translation Studies*. Amsterdam: Rodopi.  
 Ivir, V., (1987). *Procedures and strategies for the translation of culture*. in *The Indian Journal of Applied Linguistics*, Vol. 13, No. 2 (pp. 35-46), New Delhi: Bahri Publications Ltd.  
 Katterfield, E., (1997) *Uhrwerk Orange: Erziehung geschieht nicht automatisch*. Munich: Kieser Verlag.  
 Kenny, D., (2001). *Lexis and Creativity in Translation – A Corpus-based Study*. Manchester: St. Jerome Publishing.  
 Koller, W., (1992). *The Concept of Equivalence and the Object of Translation Studies in Target* (7): 2: 191-222.  
 Landers, C. E., (2001). *Literary Translation – A Practical Guide*. Clevedon: Cromwell Press.

- Mannewitz, C., (1999). *Anthony Burgess' „A Clockwork Orange“: Zur lexikalischen Systematik im Original und in Übersetzungen.* in *Interlinguistik und Lexikographie – Beiheft 5 Mitteilungsblatt der Gesellschaft für Interlinguistik e.V.* Berlin: Gesellschaft für Interlinguistik e.V.
- Shuttleworth, M., (1997). *Dictionary of Translation Studies.* Manchester: St. Jerome Publishing.
- Venuti, L & Baker, M. ed., (2000). *The Translation Studies Reader.* London & New York: Routledge.

#### Online

- Anon., (1997). *A Clockwork Orange.* Berlin: MalcolmMcDowell.net. Available at:  
URL:<http://www.malcolmmcdowell.net/>
- Anon., (2002). *Gesellschaft für Interlinguistik e. V. (GIL).* Berlin: GIL. Available at:  
URL:<http://www.interlinguistik-gil.de/>
- Anon., (2005). *A Clockwork Orange Study Guide.* Book Rags. Available at:  
URL:<http://www.bookrags.com/about/>
- Anon., (2006). *A Clockwork Orange Study Guide.* New York: Sparks Notes. Available at:  
URL:<http://www.sparknotes.com/lit/clockworkorange/index.html>
- Anon., (2007). *Calque.* St. Petersburg: Wikipedia. Available at:  
URL:<http://en.wikipedia.org/wiki/Calque>
- Anon., (2007). *Nadsat.* St. Petersburg: Wikipedia. Available at:  
URL:<http://en.wikipedia.org/wiki/Nadsat>
- Fail, L. (1999). *An Approach to the Translation of Literature: Rich Points & What They Reveal.* Available at: URL:<http://www.proz.com/translation-articles/articles/39/1/An-Approach-to-the-Translation-of-Literature:-Rich-Points-&-What-They-Reveal>
- Hyman, S. E., (1963). *Nadsat Dictionary.* Available at: URL:<http://soomka.com/nadsat.html>
- Katterfeld, E., (1997). *Uhrwerk Orange: Erziehung geschieht nicht automatisch.* Available at: URL:[http://homepages.compuserve.de/ekatterfeld/clockwork\\_Inhalt.html](http://homepages.compuserve.de/ekatterfeld/clockwork_Inhalt.html)
- Mailhac, J.P. (2007). *Formulating strategies for the translator.* Poughkeepsie, NY: Translation Journal. Available at: URL:<http://accurapid.com/journal/40strategies.html>
- Ñóglendel, N. *Wolfgang Kregge.* Cottbus: Elbenwaldforum. Available at:  
URL:[http://www.elbenwaldforum.de/printthread.php?Board=Sonstiges\\_15&main=2357203&type=post](http://www.elbenwaldforum.de/printthread.php?Board=Sonstiges_15&main=2357203&type=post)
- Token, (2006). *Grottenschlecht übersetzt!* Munich: Amazon.de. Available at:  
URL:[http://www.amazon.de/Clockwork-Orange-Anthony-Burgess/dp/3453130790/ref=sr\\_1\\_1/302-9821358-2324805?ie=UTF8&s=books&qid=1187871291&sr=1-1](http://www.amazon.de/Clockwork-Orange-Anthony-Burgess/dp/3453130790/ref=sr_1_1/302-9821358-2324805?ie=UTF8&s=books&qid=1187871291&sr=1-1)

# Importance of quantifying orthotic correction of Adolescent Idiopathic Scoliosis: A retrospective study

**Abdulla Rafiu, Stephen Hutchins**, *School of Healthcare Professions, University of Salford*

## Abstract

**Study Design.** A review of literature on the importance of quantifying orthotic correction of Adolescent Idiopathic Scoliosis (AIS).

**Objective.** To analyse the relationship and importance of parameters that are directly linked to this deformity in relation to three-dimensional matrices that are evidenced in the literature.

**Summary of background data.** Orthotic treatment for AIS has been routine practice for many years and different authors have quantified orthotic correction using different sets of parameters in one, two or three planes. Little is understood about the overall orthotic correction of AIS using measurements available in three-dimensional matrices.

**Method.** A study was carried out by reviewing literature related to orthotic treatment of AIS using the Salford University Library Databases. The search was limited to English language only.

**Results.** The literature search demonstrated that there are one or more parameters in each cardinal plane that are geometrically linked and the correction achieved in one plane may not be mirrored in other planes. No single paper was found that analysed curve correction data for all the parameters that are directly linked to AIS.

**Conclusion.** Since AIS is a three-dimensional deformity possessing geometrically-linked parameters in all three planes, tri-planar measurements need to be taken into consideration when quantifying orthotic correction of AIS.

Key words: Adolescent Idiopathic Scoliosis (AIS), three dimensional matrices, overall balance summation (OBS).

## Introduction

Adolescent Idiopathic Scoliosis (AIS) is a three-dimensional deformity of the spine of unknown origin in children aged > 10 years, which results in alteration in the geometry of functional spinal units in all three planes of motion beyond normal acceptable limits. AIS may be described as lateral spinal curvature in coronal, lordosis and kyphosis in sagittal and vertebral rotation in transverse planes beyond their normal range. There are many theories for the aetiology of this deformity which include factors within the spine (intrinsic) or outside the spine (extrinsic). However, no single cause has been found and the condition thus termed as multifactorial (Kouwenhoven, J.W. and Castelein, 2008).

The brace treatment is said to have failed if a scoliotic spinal curve increases by 5 degrees or more following onset of bracing, or if operative intervention had been needed (Lonstein and Winter, 1994; Trivedi and Thomson, 2001), if the secondary curve exceeded the primary curve by more than 5 degrees (Katz et al, 1997) or if the type of brace had been changed during the course of the treatment (Trivedi and Thomson, 2001; Wong et al, 2008). Many studies during the past ten years have demonstrated that altering the natural history of AIS by orthotic intervention is a proven fact that can no longer be questioned (Winter, 2000). Though the basic science is proven, the outcome of the end result of the brace treatment appears to have remained more or less unchanged during the past 20 years (Katz et al, 1997; Cochran and Nachemson, 1985; Wong et al, 2000; Wiley et al, 2000, Korovessis et al, 2000; Carr, 1980). The response of the curve to the brace treatment appears to be the same irrespective of the curve pattern. The pattern of correction follows three main stages; the immediate in-brace phase, the in-brace phase and the post-brace phase. In all studies examined, it was found that there is a significant decrease in Cobb's angle in immediate in brace phase, followed by further reduction in Cobb's angle during the early part of in-brace phase. The best correction is achieved during this phase; usually within the first year of orthotic treatment (Wong et al, 2000; Carr, 1980). However, there appears to be gradual loss of correction in Cobb's angle

after about one year which continues until the brace is discontinued. The final outcome remains as  $\pm 5^\circ$  of the pre-brace value.

Since statistics show that with bracing, one third of the patients show progress whereas without bracing two third of the patients progress (Winter, 2000). It therefore proves that spinal orthoses have the potential to achieve better results. We therefore, postulated that better results (i.e. curve correction) could be achieved if all the geometrically-linked parameters of AIS were considered to quantify overall orthotic correction. A literature search was therefore conducted by following the definition of scoliosis and the principles of spinal orthotic treatment to see how the parameters of correction of orthotic treatment have been quantified in previous publications.

## **Geometrical parameters in coronal plane**

### *Cobb Angle*

Measurement of Cobb angle has been the primary radiographic criteria used for the evaluation of results of treatment of scoliotic deformity for more than half a century (Majdouline et al, 2007). It is the primary parameter used to measure the amount of orthotic correction of AIS using a brace as indicated by the magnitude of the curve on coronal plane. However, the difference between successful and unsuccessful orthotic treatment of scoliosis lies not so much in the Cobb angle result; but mostly in considering the total status of deformity (Carlson, 2003). Carlson (2003) demonstrated that a  $29^\circ$  left thoraco-lumbar curve after taking an, x-ray of the immediate in-brace condition demonstrated a fairly well compensated (balanced) spine. When the same patient was followed up after 19 months, a subsequent x-ray showed the Cobb's angle reduced to 23 degrees. However, the deformity had increased in the coronal plane from 2.8cm to 3.8cm by deviating L3 vertebral segment away from the Central Sacral Line (CSL) leading to a shift the head to the left over the pelvis (known as decompensation). Nicholson, Smith et al (2003) demonstrated the same effect. A positive effect on the sagittal plane therefore produced an unwanted effect on the coronal plane. The positive effect can therefore be misleading if we forget one of the three biomechanical principles of spinal orthosis; end point control.

Wong (2000) in his study of 33 subjects with AIS showed when Cobb's angle was reduced from 34.4 degrees to 23.7degrees on the 20<sup>th</sup> month of follow up, the corresponding thoracic kyphosis had decreased from 25.9 degrees to 18.3 degrees leading to a statistically significant hypokyphosis. In contrast, Cochran and Nachemson (1985) followed 85 subjects with AIS for average of 7.5 years and found that their curves were of equal size ( $33^\circ \pm 3^\circ$ ) as when the treatment started ( $33^\circ \pm 3^\circ$ ) and none exhibited hypokyphosis ( $<20^\circ$ ). Since scoliosis is a three-dimensional deformity, the parameters that are linked to Cobb's angle have to be taken into consideration while quantifying overall correction of scoliosis. If there was a good Cobb angle correction but poor spinal balance on the coronal plane, then the correction would be unacceptable (Nicholson et al, 2003). The evidence in the literature would therefore suggest a direct link between alteration to Cobb's angle and alteration to parameters on the other two cardinal planes.

### *Decompensation / Overall Balance Summation ( OBS)*

Decompensation is the measure of lateral vertebral displacement from mid sagittal plane and it is important for two reasons: 1) it is a major factor in cosmetic appearance, and 2) the size of gravitational loads which tend to increase the magnitude of a curve are directly related to the spine's deviation from the midline (Carlson, 2003). It has been measured by some authors (Wong et al, 2008; Cochran and Nachemson, 1985; Wong et al 2000; Korovessis et al, 2000) Nachemson and Peterson, 1995), but not by others (Lonstein and Winter, 1984; Wiley et al, 2000; Pham et al, 2008; Wong et al, 2005). Decompensation does not give the magnitude of the angle formed by the scoliotic curves, but it gives the amount of balance achieved by the spine and a measure of cosmetic appearance gained. It is measured radiographically in two ways; by taking the lateral deviation of centroid of apical vertebrae from the Central Sacral Line (CSL) (Korovessis et al, 2000) or by the Overall Balance Summation (OBS) method which is by taking the distances of the centroid from the midline of all the apical vertebrae and

calculating the algebraic sum of the lateral deviations of all the curves taking right side as positive (Carlson and Cpo, 2007).

### *Sacral/ Pelvic Obliquity*

The sacrum acts as the foundation of the spine. Ignoring the pelvis and correcting the scoliotic curve would be like building a vertical wall on an inclined foundation. No matter how much we try to correct the curve there will always be an unbalanced force. The purpose of pelvic portion of the Milwaukee brace and other rigid TLSOs is the fact that it provides neutral reference to the base of the spine. No spinal orthoses can be said biomechanically effective, unless control of the end points of the curve is taken as one of the biomechanical principles of a spinal orthosis. However, no prospective study was found that included the relationship between the pelvic obliquity and orthotic correction of AIS demonstrated (Table 1). A retrospective study published by Kartz and Durrani (Katz and Durrani, 2001) found that lumbar-pelvic relationship angle significantly influenced orthotic treatment. Hoikka et al. (2001) in their study of one hundred adult patients showed strong correlation between the sacral tilt and lumbar scoliosis when the tilt was more than 3° but poor when it was smaller.

### **Geometrical parameters on the sagittal plane**

The effect of orthotic treatment on frontal plane has been well documented by several retrospective studies, but its effects on the other planes of deformity is poorly understood (Korovessis et al, 2001). Carlson (2003) described two different types of sagittal curvatures in scoliotic patients. Type I is characterised by a relatively vertical sacrum and all sagittal curvatures are reduced in magnitude. Type II is characterised by a relatively horizontal sacrum causing abnormally long lumbar lordosis extending well into the thoracic spine.

### *Thoracic kyphosis*

Thoracic hypokyphosis is cosmetically disturbing and it is one of the most important factors in thoracic distortion and volume reduction that, in extreme cases, lead to cardiopulmonary compromise. Thoracic kyphosis is mainly influenced by spinal deformity in AIS (Majdouline, 2007). Reduction of thoracic kyphotic angle is persistently demonstrated in the literature (Wong et al, 2000; Carr, 1980; Nachemson and Peterson, 1995; Chu et al, 2006). One paper demonstrated that surgery was needed to correct the deformity [2]. Others demonstrated no significant change (Cochran and Nachemson, 1985; Carr, 1980), whilst others demonstrated that, by modifying the brace biomechanically in a more effective way, it was able to bring the thoracic hypokyphosis to a more normal configuration (Carlson, 2003). Normal thoracic kyphotic angle is between 20°- 40° (Carr, 1980; Harvey et al, 2002).

### *Lumbar Lordosis and Pelvic Incidence*

Lumbar lordosis and pelvic incidence are two distinct parameters in the sagittal plane but are closely related. Pelvic incidence and sacral slope strongly determine lumbar lordosis (Mac-Thiong et al, 2003) but no significant differences have been found between the pelvic incidence in normal subjects and those with scoliosis (Majdouline et al, 2007) However, more recent studies done by Mac-Thiong et al. (2003), in their retrospective study of 160 patients with AIS, found that pelvic incidence in AIS was higher than that of normal adolescents but that it did not associate with curve types. They also found that lumbar lordosis was influenced mainly by the sagittal alignment of pelvis in AIS. This agrees with the study done by Upasani et al. (2007) in their retrospective chart review of 53 AIS patients along with a control group comprising of 50 normal adolescents. The pelvic incidence appears to be the main axis of the sagittal balance of the spine and it is the key factor for managing spinal balance (Legaye et al, 1998).

### **Geometrical parameters in transverse plane**

### *Apical Vertebral rotation*

Axial rotation of the apical vertebra is important for the pathomechanics of the 3D deformity in scoliosis (Majdouline et al, 2007) and vertebral de-rotation reduced by bracing significantly

improves cosmesis in AIS (Korovessis et al, 2000). Upadhyay et al. (1995) in his study of eighty five patients with AIS showed increases in vertebral rotation and /or in Cobb's angle after brace application. Curve progression demonstrated that the brace failed to give adequate curve correction or slowing down of curve progression in 93% of patients and 79% of these subsequently underwent surgery. However, the patients with no change in both vertebral rotation and Cobb's angle after bracing also often experienced brace failure (69%); two of which (15%) required surgery. This indicated that there is a strong relationship between Cobb's angle and vertebral rotation and increase in one or both can lead to failure of the brace in giving adequate correction.

## Discussion

AIS is a three dimensional deformity in nature and the deformity lies on all three planes, namely coronal, sagittal and transverse. Each plane has well defined parameters that have a predominant effect on scoliosis and they can be linked both within the same plane and also to other planes. These parameters are geometrically linked and can have positive and negative effects when manipulated on one plane. Therefore ignoring one parameter while calculating orthotic correction could lead to sub-optimal treatment.

Table 1 shows a summary of how orthotic correction is measured by the ten most relevant prospective studies found in the literature. Cobb's angle was measured in all the studies. However, no studies included pelvic parameters on coronal or sagittal plane irrespective of end point control being one of the principles of spinal orthotics. Different braces used in the studies cannot be compared; however, overall end results of the treatments remained  $\pm 5^\circ$  from the pre-brace value.

We therefore postulate from this study, that the reason for this pattern remaining unchanged over the years of orthotic treatment is the exclusion of pelvic parameters in the studies examined. We further hypothesise that by considering pelvic parameters on both coronal and sagittal planes, and acting on them, a better overall outcome would result because pelvic orientation is the key factor for maintaining the spinal balance on sagittal and coronal planes.

**Table1.** Summary of literature review on quantification of orthotic correction in AIS in prospective studies.

Study	Year	Brace	Subjects	Geometric Parameters						
				Coronal Plane			Sagittal Plane			Transverse Plane
				Cobb's Angle	Decom	PO	TK	LL	PI	AVR
J.E.Lonstein & R.B.Winter[2]	1994	Milwaukee	1020	x	-	-	x	-	-	x
Nachemson et al[15].	1995	TLSO	111	x	x	-	x	-	-	x
Cochran and Nechemson[7].	1983	Milwaukee	85	x	-	-	-	-	-	-
Wiley et al[29].	2000	Boston Brace	50	x	-	-	-	-	-	-
Korovessis et al[10].	2000	TLSO	24	x	x	-	x	x	-	x
Wong et al[8].	2000	TLSO	26	x	x	-	x	x	-	x
Wong et	2005	CAD-CAM	40	x	-	-	-	-	-	x



al[30].		+Manual TLSO								
Chu et al[23].	2006	TLSO	14	x	-	-	x	-	-	x
Pham et al[16].	2008	Cheneau brace	32*	x	-	-	-	-	-	-
Man Sang Wong et al[31].	2008	SpineCore +Rigid TLSO	43	x	x	-	-	-	-	x

\*Idiopathic Scoliosis, Decom =Decompensation, PO=Pelvic Obliquity, TK=Thoracic Kyphosis, LL=Lumbar Lordosis, PI=Pelvic Index, AVR=Apical Vertebral Rotation

Figure 1 shows a model to obtain quantification of AIS by considering the definition of scoliosis and principles of orthotic treatment. Significant parameters of three-dimensional deformity obtained from the review of previous studies have been included. These tri-planar geometrically-linked parameters were obtained as inputs to curve correction achieved by orthotic intervention via CTLSO or TLSO application. By applying orthotic principles in addition to curve correction, endpoint control and balanced horizontal forces may be taken into consideration. Endpoint control includes biomechanically controlling the two ends of the spine i.e head and pelvis, which involves attaining balance of head over the pelvis, (as measured in OBS) and more importantly, providing a neutral reference to the spine by calculating pelvic obliquity and the pelvic index. The third principle, known as balanced horizontal forces, is provided by the pads within the spinal orthoses and force can be calculated using pressure sensors.

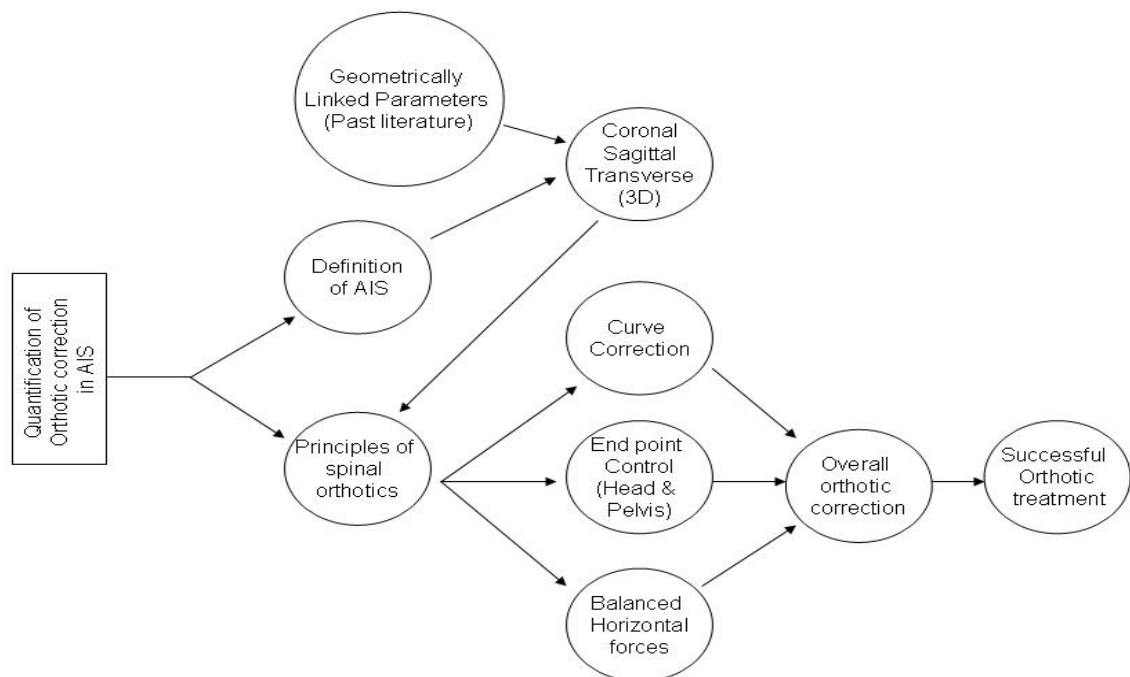


Figure 1. Quantification of orthotic correction in AIS with reference to definition of scoliosis and principles of spinal orthotics.

### Key points

- AIS is a three-dimensional deformity comprising geometrically-linked parameters in all three planes of motion;
- Investigations into the efficacy of orthotic intervention in the form of spinal bracing and the design of such orthoses should involve consideration of correction in all three planes;

- The pelvis is the foundation of spine and has a significant influence on the spine in both coronal and sagittal planes of motion;
- Orthotic correction should be quantified and represented as a vector quantity as the deformity has magnitude and direction in all three planes of motion;
- There is a need for a global formula for orthotic correction of AIS.

## References

- Carlson, J.M. and F. Cpo, Treatment of Adolescent Idiopathic Scoliosis. *JPO* 2007. 19(3): 64-66.
- Carlson, J.M., Clinical Biomechanics of Orthotic Treatment of Idiopathic Scoliosis. *JPO*, 2003. 12(4[Supl.]): S17-S30.
- Carlson, J.M., Clinical Biomechanics of Orthotic Treatment of Idiopathic Scoliosis. *JPO*, 2003. 15(4): S17-S30.
- Carr, W.A., Treatment of idiopathic scoliosis in the Milwaukee brace: long-term results. *The Journal of Bone and Joint Surgery*, 1980. 62(4): 599-612.
- Chu, W.C.W., Wong, M. S., Chau, W. W., Lam, T. P., Ng, K. W., Lam, W. W. M., et al., Curve correction effect of rigid spinal orthosis in different recumbent positions in adolescent idiopathic scoliosis (AIS): A pilot MRI study. *Prosthetics and orthotics international*, 2006. 30(2): 136-144.
- Cochran, T., and Nachemson, A., Long-term anatomic and functional changes in patients with adolescent idiopathic scoliosis treated with the Milwaukee brace. *Spine*, 1985. 10(2): 127-133.
- Harvey, Z., Chamis, M. and Lin, R., The Impact of Lateral Pads versus Posterolateral Pads in the Management of Idiopathic Scoliosis. *JPO*, 2002. 14(4): 165-169.
- Hoikka, V., M. Ylikoski, and K. Tallroth, Leg-length inequality has poor correlation with lumbar scoliosis. A radiological study of 100 patients with chronic low-back pain. *Archives of orthopaedic and trauma surgery*, 1989. 108(3):173-175.
- Katz, D.E. and A.A. Durrani, Factors that influence outcome in bracing large curves in patients with adolescent idiopathic scoliosis. *Spine*, 2001. 26(21): 2354-61.
- Katz, D.E., Richards, B. S., Browne, R. H., Herring, J. A., A comparison between the Boston brace and the Charleston bending brace in adolescent idiopathic scoliosis. *Spine*, 1997. 22(12): 1302-1312.
- Korovessis, P., Kyrkos, C., Piperos, G., Soucacos, P.N., Effects of thoracolumbosacral orthosis on spinal deformities, trunk asymmetry, and frontal lower rib cage in adolescent idiopathic scoliosis. *Spine*, 2000. 25(16): 2064-2071.
- Kouwenhoven, J.W. and Castelein, R.M., The pathogenesis of adolescent idiopathic scoliosis: review of the literature. *Spine*, 2008. 33(26): 2898-2908.
- Legaye, J., Duval-Beaupere, G., Hecquet, J., Marty, C., Pelvic incidence: a fundamental pelvic parameter for three-dimensional regulation of spinal sagittal curves. *Eur Spine J*, 1998. 7(2): 99-103.
- Lonstein, J.E. and Winter, R.B., The Milwaukee brace for the treatment of adolescent idiopathic scoliosis. A review of one thousand and twenty patients. *JBJS*.1994: 1207-1221.
- Mac-Thiong, J.M., Labelle, H., Charlebois, M., Huot, M. P., de Guise, J. A., Sagittal plane analysis of the spine and pelvis in adolescent idiopathic scoliosis according to the coronal curve type. *Spine*, 2003. 28(13): 1404-1409.
- Majdoulina, Y., Aubin, C. E., Robitaille, M., Sarwark, J. F., Labelle, H., Scoliosis correction objectives in adolescent idiopathic scoliosis. *J Pediatr Orthop*, 2007. 27(7): 775-781.
- Nachemson, A.L. and L.E. Peterson, L.E., Effectiveness of treatment with a brace in girls who have adolescent idiopathic scoliosis. A prospective, controlled study based on data from the Brace Study of the Scoliosis Research Society. *JBJS*. 1995: 815-822.

- Nicholson, G.P., Ferguson-Pell, M. W., Smith, K., Edgar, M., Morley, T., The objective measurement of spinal orthosis use for the treatment of adolescent idiopathic scoliosis. *Spine*, 2003. 28(19): 2243-2251.
- Pham, V.M., Houilliez, A., Schill, A., Carpentier, A., Herbaux, B., Thevenon, A., Study of the pressures applied by a Chêneau brace for correction of adolescent idiopathic scoliosis. *Prosthetics and orthotics international*, 2008. 32(3): 345-355.
- Trivedi, J.M.O., M. Ch, and Thomson, J., Results of Charleston Bracing in Skeletally Immature Patients With Idiopathic Scoliosis. *Journal of Pediatric Orthopaedics*, 2001. 21(3): 277-280.
- Upadhyay, S.S. and Nelson, I.W., New prognostic factors to predict the final outcome of brace treatment in adolescent idiopathic scoliosis. *Spine*, 1995. 20(5): 537-545.
- Upasani, V.V., Tis, J., Bastrom, T., Pawelek, J., Marks, M., Lonner, B. et al., Analysis of Sagittal Alignment in Thoracic and Thoracolumbar Curves in Adolescent Idiopathic Scoliosis: How Do These Two Curve Types Differ? *Spine*, 2007. 32(12): 1355-1359.
- Wiley, J.W., Thomson, J. D., Mitchell, T. M., Smith, B. G., Banta, J. V., Effectiveness of the Boston brace in treatment of large curves in adolescent idiopathic scoliosis. *Spine*, 2000.25(18): 2326-2332.
- Winter, R.B., Expert Editorial: Bracing for Idiopathic scoliosis: Where do we go now? *JPO*, 2000, 12(1): 2-4.
- Wong, M.S., Cheng, J.C., Lam, T.P., Ng, B.K., Sin, S.W., Lee-Shum, S.L., et, al., The Effect of Rigid Versus Flexible Spinal Orthosis on the Clinical Efficacy and Acceptance of the Patients With Adolescent Idiopathic Scoliosis. *Spine*, 2008. 33(12): 1360-1365.
- Wong, M.S., J.C.Y. Cheng, and K.H. Lo, A comparison of treatment effectiveness between the CAD/CAM method and the manual method for managing adolescent idiopathic scoliosis. *Prosthetics and orthotics international*, 2005. 29(1): 105-111.
- Wong, M.S., Mak, A. F. T., Luk, K. D. K., Evans, J. H., Brown, B., Effectiveness and biomechanics of spinal orthoses in the treatment of adolescent idiopathic scoliosis (AIS). *Prosthetics and Orthotics International*, 2000. 24(2): 148-162.

# A Framework for Geo-Spatial Data Sharing

**Norzamni Salleh and Professor Farzad Khosrowshahi**, *School of the Built Environment, University of Salford*

## Abstract

Geo-spatial data are items of information related to a location on earth. Accurate, up-to-date, relevant and accessible geo-spatial data is becoming essential input in supporting many operational and strategic decisions for many disciplines. With increasing demand in the use of geo-spatial data by various organizations, there is also an increasing demand for shared and easy access to these geo-spatial data. This is further aggravated by the easy availability of internet connectivity and the advancement of web technology.

However, there are technical and non-technical issues which are major hindrances in sharing these geo-spatial data. Among them are interoperability, operational policy, data models, legal issues, standards and etc. Hence there is a necessity to build a framework which addresses these issues in order to achieve a seamless geo-spatial data across organizations.

In this paper, Brunei, a small country in South East Asia, was chosen as a case study area. A preliminary guideline on factors affecting geo-spatial data sharing has been developed entirely from the existing literature and used as a basis for data collection. The data collected from case studies were used to refine the preliminary guidelines to include factors that affect geo-spatial data sharing in Brunei.

Finally, based on this preliminary guideline, a conceptual framework for Brunei geo-spatial data sharing was customized. The framework needs to be validated to ensure the effectiveness.

**Keywords** : geo-spatial data, geo-spatial data sharing, framework

## 1.0 Introduction

### 1.1 *What is geo-spatial data?*

Geo-spatial data is information that is defined spatially (in location) by four dimensions (geometry and time) related to the earth (Grooth & McLaughlin, 2003). There were various other terms used in the literature such as geographic data, geographic information, geo-spatial information, spatial data and spatial information (Masser, 1998; Groot & McLaughlin, 2000). Since this research focused on base mapping data i.e. raw data, the term geo-spatial data will be used throughout this paper.

With the advance of technology, decision makers of various disciplines rely on accurate, up-to-date, relevant and accessible geo-spatial data in supporting their operational and strategic decisions. This has increased the demand for easy access to geo-spatial data and has pushed multi-organizations to share, as rarely do all these geo-spatial datasets reside within one organization.

### 1.2 *Why share geo-spatial data?*

There were many benefits for multi-participants engaging in geo-spatial data sharing found in the literatures (National Research Council, 1993; Nedovic-Budic, 1994; Azad & Wiggins, 1995; Onsrud & Rushton, 1995; Kevany, 1995; I.Williamson et al. 2003; Ian Masser, 2005) which include:-

- Reducing costs of data production and eliminating duplication of effort
- Reducing time spent in collecting data
- Inclusion of more diverse maps in decision making
- Developing applications more quickly and easily by using existing data and data development standards
- Providing better data for decision-making

- Saving development effort by using fundamental and standardized data guidelines and tools.
- Ability to perform cross-jurisdictional and cross-sectoral decision making, analysis and operations based on common data and understanding of issues
- Expanding market potential through recognition and credibility as an SDI participant as well as through the formation of beneficial partnership
- Providing consolidated directions to vendors regarding required technical features
- Facilitating the development of knowledge infrastructure and communication networks.
- Providing efficient services
- Increasing the value and quality of information
- Improved morale and learning self-confidence was highlighted as an intangible benefit (Nedovic-Budic & Pinto, 1999).

Despite all these benefits, geo-spatial data sharing remain an issue in most countries (Nedovic-Budic & Pinto, 2001; Warnest, 2005). There were various factors that can be categorized as technical and non-technical as well as institutional factors that impede geo-spatial data sharing (Onsrud & Rushton, 1992). The technical factors are mainly related to problems in coordinating system requirements (Croswell, 1991; Calkins & Weatherbe, 1995); lack of common data definition, formats, and models (Frank, 1992; Dawes 1996); differences in data quality (Frank, 1992) and networking costs (Nedovic-Budic & Jeffrey Pinto, 1999).

However, technical factors are well studied and have been largely resolved, while non-technical aspects are less well studied but equally important (Campbell & Masser, 1995). This is further supported by Budic & Pinto (1999); resistance to data sharing is not normally related to technical issues, that is, incompatible systems or data structures, but the challenges are first and foremost 'people' challenges. The author further emphasized 'people represent the need to better isolate and address the human factors that are likely to impede free data sharing across organizational boundaries'.

Non-technical factors include data confidentiality, liability, pricing (Campbell & Masser, 1995), lack of negotiation, institutional inertia (Craig, 1995), fear of losing autonomy over control of information and organizational power (Pinto & Azad, 1994; Azad & Wiggins, 1995; Meredith, 1995), legal and public policy (Onsrud & Rushton, 1992). Most of these issues were due to the variations in priorities between participants, the differences in the ability to exploit GIS facilities, the differences in the level of awareness and spatial data handling skills and agreements over access to information, leadership, data standards, equipment and training (Campbell & Masser, 1995). Williamson (2003), criticized the fact that the issues of cost recovery, copyright and legal liability have done little to encourage organizations to provide access to spatial information.

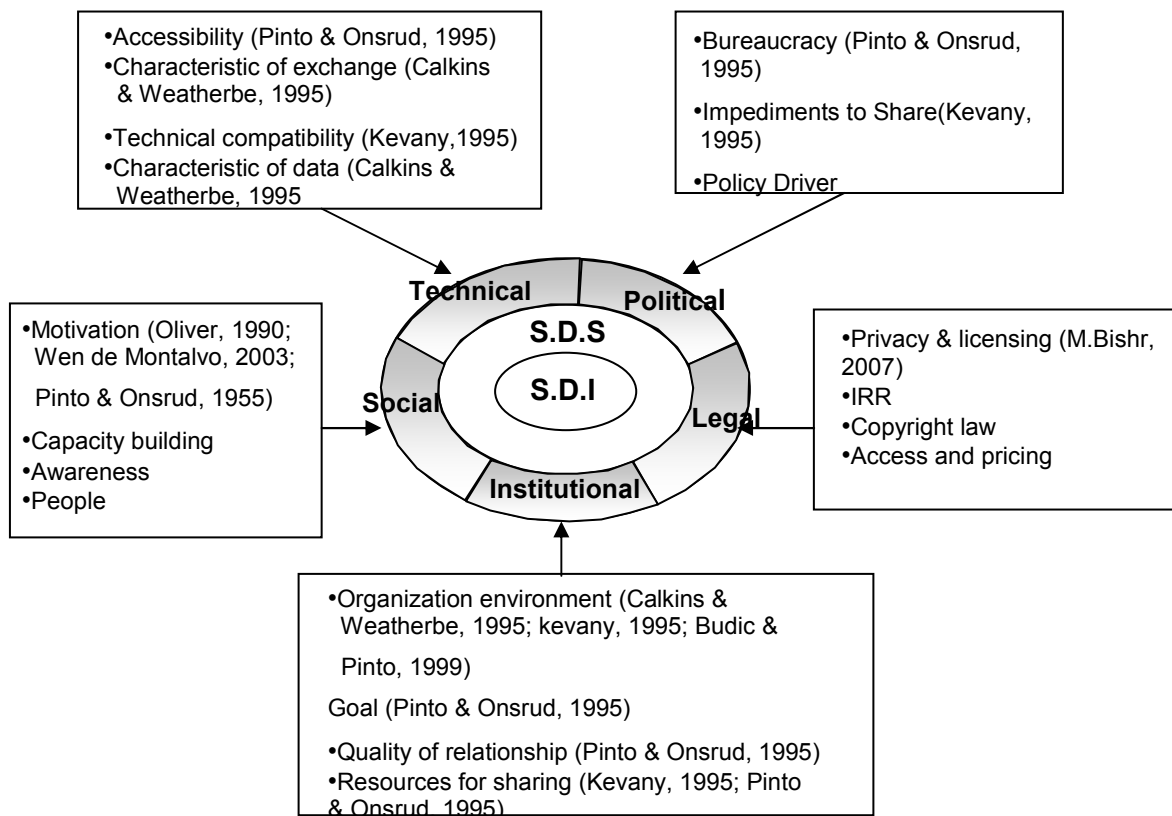
### *1.3 Existing theory data sharing frameworks*

There were various theory frameworks on data sharing found in the literature. Among them are antecedents and consequences of information sharing (Pinto & Onsrud, 1995), taxonomy for research into spatial data sharing (Calkins & Weatherbe, 1995), factors relevant to GIS data sharing (Kevany, 1995), typology of inter-organizational relations and dynamics (Azad & Wiggins, 1995), a typology of 6 determinants of inter-organizational relationships (Oliver, 1990), understanding organizational data sharing framework (Nedovic-Budic & Pinto, 1999), a model of willingness to share data in South Africa (When de Montalvo, 2003), interaction between organizational behavior of spatial data sharing and social and cultural aspects (Omran et al. 2006), a model for national spatial data infrastructure collaboration (Warnest, 2005) and the data sharing partnership model (McDougall, 2006). Most of these frameworks were based on the authors' experiences and have not been proven empirically except for Nedovic-Budic & Pinto's (1999), When de Montalvo's (2003) and McDougall's (2006).

### *1.4 Development of preliminary framework*

All the existing theory frameworks together with the literature review emphasized the importance of developing an improved understanding level of motivations and barriers for organizational cooperation. Based on these findings and understanding, a preliminary

guideline on factors affecting geo-spatial data sharing (Figure 1.1) was designed and used as an aid for data collection.



**Fig 1.1:** Preliminary Guideline on Factors Affecting Geo-spatial Data Sharing

#### 1.4 Brunei as a case study area

Brunei was chosen as a case study area because the country is at the early stage of moving towards Spatial Data Infrastructure (SDI). Various government departments gather their own geo-spatial data specifically for internal use which has resulted in fragmented datasets. With about 85% of the country's data were in digital, it is the right time to ensure that data collected were not duplicated and can be integrated as well as shared among users to support the implementation of SDI in Brunei.

## 2.0 Methodology

### 2.1 Research methodology and philosophy

Research methodology can be defined as the whole approach of the research process which includes the research philosophy, research strategy and research method. Research philosophy provides a guiding contact for the research approach and the research technique for successful researchers (Kagioglou et al.1998). This is supported by Easterby-Smith et al. (2002) who assert that the failure to think through philosophical issues, while not necessarily fatal, can seriously affect the quality of research, which is central to the notion of the research design. Therefore, it is vital to understand where the research stands philosophically, because it can influence the selection of the appropriate research approach.

The nature of this research requires the researcher to interact with the participant in order to come to an in-depth understanding of the real life situation and to increase general understanding of the situation by examining the perception of human actors. Table (2.1) shows a summary of the research stance philosophy and the adopted research strategy, design and method.

**Table 2.1.** Summary of the Research Methodology adopted for this research

<b>Elements of Research</b>	<b>Views</b>	<b>My Choice</b>
Ontology [What is knowledge?]	<u>Realism</u> : [Objective and external] <u>Idealism</u> : [Understood by examining]	<u>Towards Idealism</u> [To investigate real life situations in order to increase understanding of the situation]
Epistemology [How do we know knowledge?]	<u>Positivism</u> : [Observer is independent] <u>Socio-Constructivist</u> : [Observer is part of what is being observed]	<u>Towards Social-Constructivist</u> [Researcher will be actively involved in what is being observed]
Axiology [What value goes into it?]	<u>Value Free</u> <u>Value Laden</u>	<u>Towards Value Laden</u> [The value will depend on human beliefs and experience]
Approach	<u>Deductive</u> : [Theory to Empirical] <u>Inductive</u> : [Empirical to Theory]	<u>Both Deductive and Inductive</u> [Initial framework was deducted entirely from Literature Review] [Findings from case studies and knowledge gain from literature review were use to customize the conceptual framework]
Strategy  Design	<u>Experiment / Survey / Case Study / Action Research / Ethnography</u>  <u>Single Holistic / Single Embedded / Multiple Holistic / Multiple Embedded</u>	<u>Case Study</u> [Researcher requires in depth study and will need to interact with the problem environment]  <u>Multiple holistic</u> [Researcher needs to enquire form more than one department to gain full view of the reality]
Method	<u>Literature Review / Interview / Documentation / Observation / Questionnaire</u>  <u>Qualitative / Quantitative / Mixed Method</u>	<u>Literature Review / Semi-structured Interview / Focis Group meeting / Supporting documents</u>  <u>Qualitative (descriptive)</u> [The research traces the sequence of events over time and describes the key phenomena]

### 3.0 Case studies

Data were gathered via semi-structured interviews, focus group meetings and through documentation. Three government departments were selected based on criteria of geo-spatial data producer and type of data users. The departments were the Survey Department, Road Department and Town & Country Planning Department. Three different sections of each selected department were focused on, which included administration, IT and applications.

Semi-structured interviews were conducted at both management and operational levels, which provided perspectives from the staff involved in the initiation, design, implementation, management and operational process that relates to geo-spatial data sharing. A total of 12 officers, which comprised 5 males and 7 females from the three selected departments, were interviewed, ranging from Head of Department, engineers, IT experts, GIS specialist, and senior planners and mapping officers. This will enable the researcher to explore the organizational structure of work more closely.

Focus groups were also arranged on two occasions; one was with a group of users that work daily with the data and the other group was the data providers' group. Both groups were composed of technician level staff. The aim of these focus group meetings was to get views from both data providers and users on their experience in accessing and maintenance of the available geo-spatial data.

Documents relevant to the research interest were also compiled to gain a clear understanding of the case being studied, and to support what had not been covered during the interviews and the focus meetings.

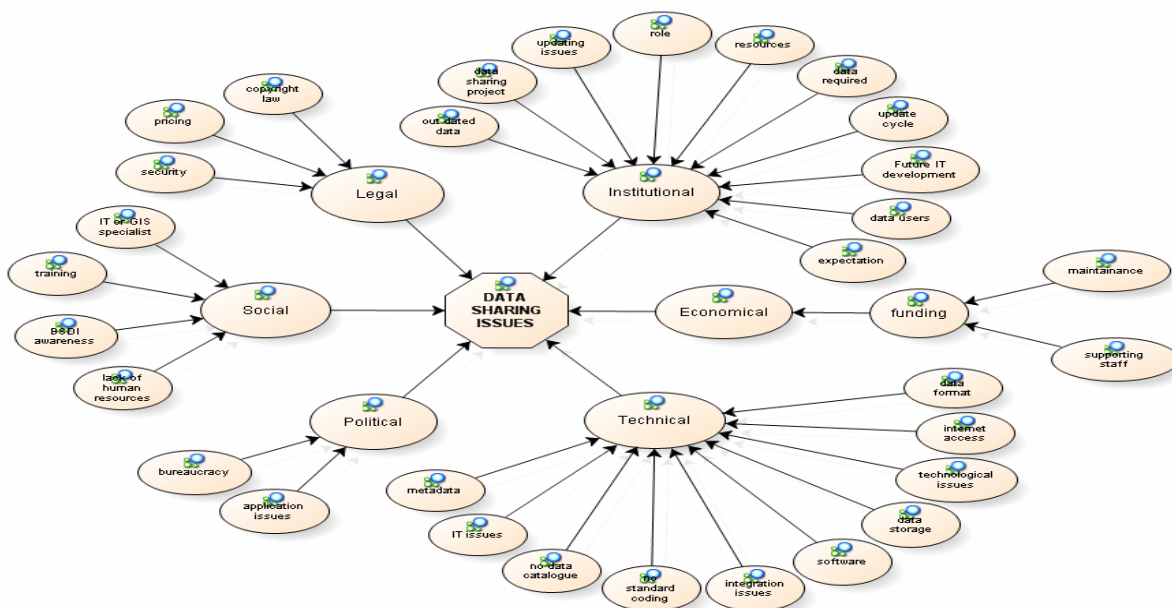
#### 3.1 Data analysis

Qualitative research software from QSR International, known as NVivo, was used to help with the analysis and representation technique for mapping interaction and relationships between the key factors of the research.

In order to construct the cognitive maps, individual interview documents were transferred and stored as a case in the NVivo software. Then relevant statements in every case were assigned codes according to a group of factors.

Having been coded and assigned to groups, the data were analyzed and the relationships between nodes linked. The output displayed the nodes clustered in a hierarchical order linked to the key concept using lines and arrows used to identify the meaning between them and concepts.

The diagram (Figure 3.1) explains that there were six [6] group of factors affecting geo-spatial data sharing in Brunei. Listed in (Table 3.1) is a compilation of current issues faced by government departments in Brunei.



**Figure 3.1.** Factors Affecting Geo-Spatial Data Sharing (case studies result)



**Table 3.1.** Compilation of current issues in utilizing spatial data across the 3 case studies

- Data accuracy and currency
- Duplication of data
- Slow and inconvenient, frustrating to access data
- Limited GIS licenses for mapping systems (budget)
- Incomplete data (high cost)
- Limited high end hardware
- Poorly resourced GIS or IT staff to manage data and systems
- Data was designed for specific departmental use only
- Systems used not user friendly
- Incompatibility with other departments systems
- Lack of awareness of what data is available
- Lack of knowledge and skills of how to access data
- Lack of GIS / IT specialist or professionals
- Lack of staff with GIS capabilities and experience
- Having to go to places to obtain information
- Time required to obtain permission to use information (bureaucracy)
- Request must be made on every different project
- No data policy in place
- Custodianship of data
- Limited access to data due to privacy
- Data pricing and licensing
- No standard procedure for requesting to use data
- Difficulty in obtaining data (no data catalogue)
- Copyright law
- Department procedure (flow of work)
- Resources in different format
- Silo culture

### 3.2 Summary of the case studies' findings

From the case studies, the drawbacks to data sharing can be judged as occurring more from the non-technical perspective than the technical. Although the two were inter-related, the case studies' findings proved that most issues raised are related to political, legal, organizational, social and economic factors.

Data pricing appeared to be the major issue faced by most departments. The issue of data access, which includes bureaucracy, confidentiality and copyright law, further complicates the process. There was no policy in place that addresses these issues, thus the situation was an extremely difficult one in which to encourage sharing.

The 'silo culture' was still heavily practiced in the country. It was noted that the current vision of the different government department focuses on individual plans and had led an individual department requested budget allocations according to their own plans. The approach needs to be changed and the departments need to collaborate to enable data sharing among departments to become a reality.

There was no information on data availability, no knowledge about how to access data and no knowledge about data standards. This led individual departments to capture their own data to their own specifications which contributed to duplication and fragmented data. It is very important to implement standard and metadata to improve interoperability and efficiency in accessing data, and to enable automated data exchange.

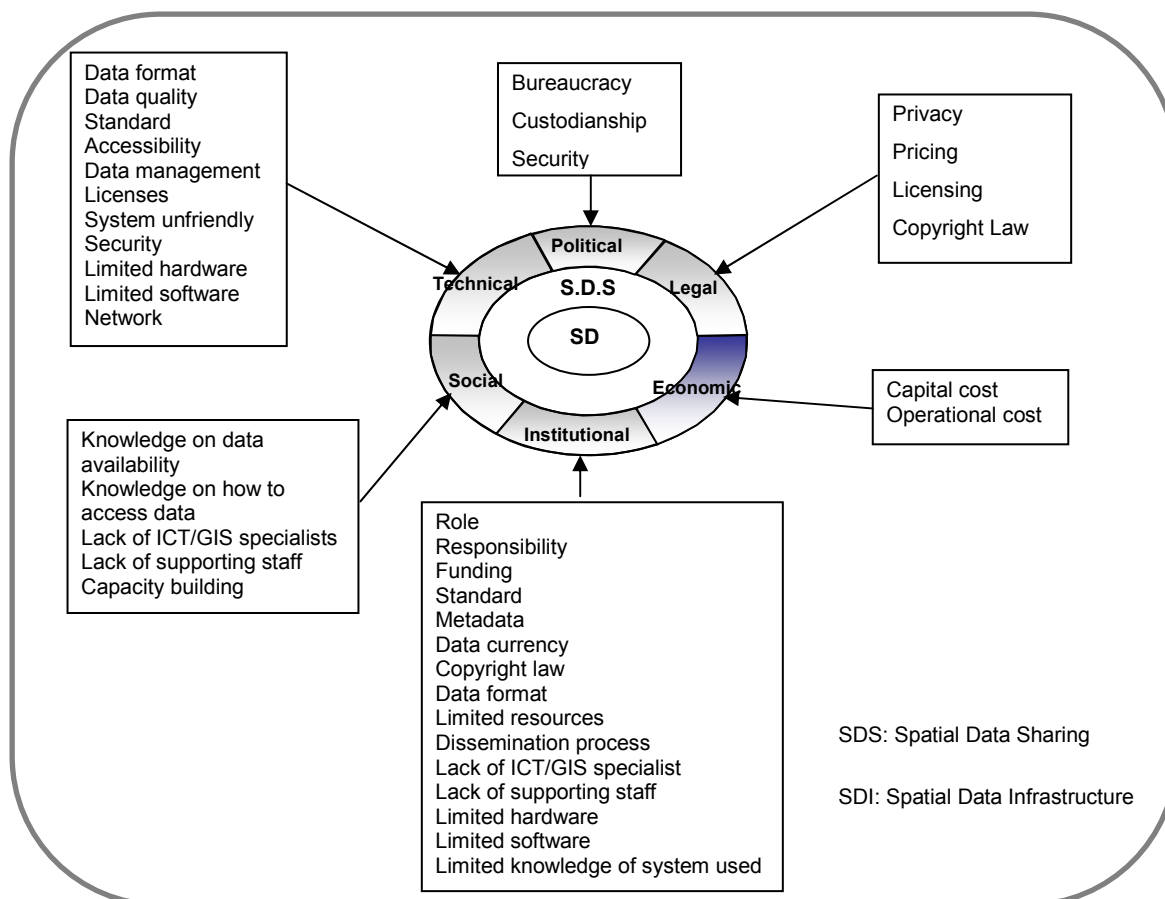
It was also noticed that the IT participants interviewed do not establish interaction with other departmental staff, as requests for maps or other information are usually initiated by different sections of the department. The key interactions between staff were predominantly one position level up or down from the respective officer. This has contributed to the lack of awareness of external issues and constraints within the sampled area.

Continuous funding and capacity building were among the key barriers to geo-spatial data sharing. Continuous funding is essential to support the maintenance of long term geo-spatial data sharing. Capacity building plays an important role in promoting and motivating geo-

spatial data sharing. At the time, there was a lack of IT/GIS specialist and supporting staff, which limited the usage of data and the technology.

Overall, the participants' responses revealed that a need to have a seamless database is essential and would contribute to providing an effective and efficient service.

Based on these findings, the preliminary guideline was modified accordingly. The economic factor was treated as an independent factor, thus the preliminary guideline comprised six different factors (Figure 3.2).

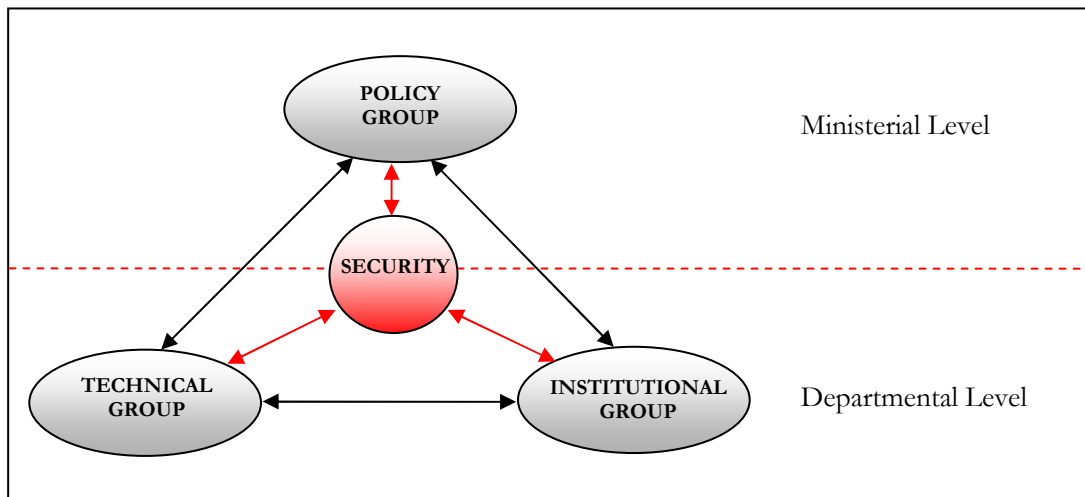


**Figure 3.2.** A Preliminary Guideline of Factors Affecting Geo-Spatial Data Sharing within Government Departments in Brunei

#### 4.0 Development of a conceptual framework for geo-spatial data sharing

Figure 3.2 showed how issues of geo-spatial data sharing within government departments in Brunei were grouped into 6 factors. These 6 factors were further classified into 3 different groups according to level of responsibilities and jurisdictions (Figure 4.1). The three groups are Policy Group (Ministerial level), Institutional Group (Departmental level) and Technical Group (IT specialist).

From the beginning of this research, it has been noted that geo-spatial data sharing is hindered by technical and non-technical factors. The non-technical factors include political, legal, institutional and social ones. Funding issues were normally addressed under individual departments, and termed as departmental budgets or projects based under the National Development Plan.



**Figure 4.1** Three groups responsible for geo-spatial data sharing

The key finding in this research with regard to geo-spatial data sharing is that the economic factor, which comprises capital cost and operational cost, should be treated as an independent factor. The 'silo culture' was influenced by the vision of individual departments whereby each department applied for its own budget according to its own plan to achieve its own vision. This culture has to be changed by introducing a vision that is acceptable for a wide environment and also by introducing central funding. This funding should be under the management of a higher authority such as Ministerial level.

Most of the literature highlighted that political and legal factors can result in policies (Budic, 1999; Mc Laughlin et al. 2003; Warnest, 2005; Mc Dougall 2006). The policies will provide clear rules and responsibilities that address issues that relate to political and legal factors. These policies will be looked after by a group of higher authority, that is responsible for policy making. In Brunei's case, policy making is at Ministerial level.

The case studies also revealed that the development of IT and uses of geo-spatial data in different government departments varies from infancy to an advanced level. This was due to the individual department's priorities. In most government departments there was a lack of IT or GIS specialist and supporting staff; this limited the usage of data and the technology. Recruiting new staff is beyond the individual department's jurisdiction and furthermore, training requires funding. In figure 3.2, the issues of capacity building, lack of specialist, supporting staff and training were categorized under social factor. This factor is dependent on the economic factor.

From the above explanations, policy which addressed political and legal factors, the economic factor which includes capital and operational cost, and the social factor which includes human resource and capacity building, should all be under the responsibility of the policy making level as stipulated in most of the literature (Budic, 1999; Williamson et al, 2003; Warnest, 2005; Mc Dougall, 2006). In Brunei's case that is the ministerial level.

The Institutional Group comprised all potential participants in the project. As Obermeyer (1995) infers, essential to collaborations is the leadership of one or more champions. This champion(s) will recommend the course of action to the policy level, including design, implementation and operation of the innovation, and will act as innovator over a long term. The author further recommends that the champion(s) should have a high level of authority in order to make and implement decisions and to take professional and personal responsibilities for the success or failure of the project. Noting this statement, the Institutional group was divided into two [2] sub-groups i.e. a management group and a working group. The Management group was composed of high level staff such as the Head of Department from each participating department.

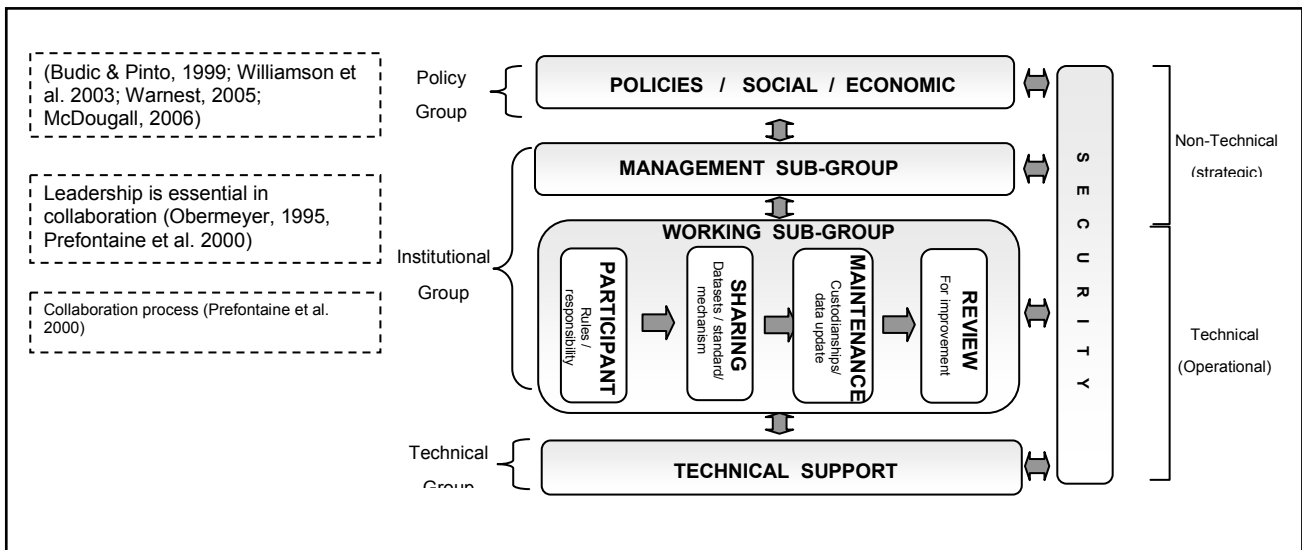
The Working group will comprise of all participating departments that need to collaborate in order to share the data in a seamless database. Prefontaine et al. (2000) highlighted four [4] components in collaboration, being negotiation, commitment, execution and assessment. These components of collaboration were used as guidelines to be customized for the Brunei

scenario.

The third group is the Technical group. This group is responsible for recommending technical strategies, framing policy needs relating to data access and data standards. In most government departments, IT specialists are lacking and departments were dependent on IT consultants with regard to IT related activities

The data sharing projects are involved with real data and therefore privacy and security is highly regarded. In Brunei, the security is governed by an independent department. Its responsibility is to advise on security matters, and its involvement in the project will start from the policy making, throughout the whole collaboration process and also in the technical structure.

Based on these findings, a conceptual framework was designed for Brunei geo-spatial data sharing, which showed the flow of work among the 3 main groups (Figure 4.2). The framework needs further refinement by breaking down the steps, and will follow up with validation.



**Figure 4.2.** A Conceptual Framework for Brunei Geo-Spatial Data Sharing

## 5.0 Conclusion

The paper addressed the problem associated with issues to spatial data sharing. The methodology which was adopted consisted of semi-structured interview, focus group meeting and documentation has provided an in-depth investigation in the 3 selected government departments in Brunei. In addition to this investigation, the existing literature has provided a solid foundation to conceptualise the framework.

The case studies' revealed that there are 6 factors that remain as barriers to spatial data sharing which includes Political, Legal, Social, Institutional, Economic and Technical. Each of these factors needs to be considered in achieving successful spatial data sharing. These findings together with the literature review, a conceptual framework for Brunei geo-spatial data sharing was customized (figure 4.1).

However, the conceptual framework still needs to be validated to ensure its effectiveness. It is important to note that this framework was derived from case studies in Brunei. Therefore, there tend to be additional or fewer issues in other countries because of difference in culture and administration.

## References

- Azad, B. & Wiggins, L. (1995) Dynamic of inter-organizational data sharing, New Brunswick, Centre for Urban Policy Research.
- Benbasat (1987) the Case Research Strategy in Studies of Information Systems. *MIS Quarterly*, 11, 369-386.
- Budhathoki, N. R. & Nedovic-Budic, Z. (2007) *Research and Theory in Advancing Spatial Data Infrastructure Concepts*, California, ESRI Press.
- Calkins, H. & Weatherbe, R. (1995) *Taxonomy of Spatial Data Sharing*, New Brunswick, Centre for Urban Policy Research.
- Campbell & I.Masser (1995) *Information Sharing: The effects of GIS on British Local Government*, New Brunswick, Centre for Urban Policy Research.
- Coleman, D., McLaghlin, J. & Nicholas, S. (1997) *Building a Spatial Data Infrastructure*. 64th Permanent Congress Meeting of the Federation International of Geomatics (FIG), Singapore.
- Collils, J. & Hussey, R. (2003) *Business Research: A practical guide for undergraduate and postgraduate students*, New York, Palgrave Macmillan.
- Committee, M. S. (1993) *Towards a Coordinated Spatial Data Infrastructure for the Nation* Washington, National Academy Press.
- Council, N. R. (1994) *Promoting National Spatial Infrastrucutre through partnerships*, Washington, National Academy Press.
- Crompvoets, J. (2007) *National Spatial Data Clearinghouses, 2000-2005*, California, ESRI Press.
- Easterby-Smith, M., Thorpe, R. & Lowe, A. (2004) *Management Research: An Introduction*, California, Sage Publication Ltd.
- Eurogi (2005) Eurogi website : Rationale, Missions and Objectives of EUROGI.
- Graglia & Annoni(2007) *INSPIRE : An innovative approach to the development of SDI in Europe*, Redlands, ESRI Press.
- Graglia, M. & Annoni, A. (2007) *INSPIRE : An Innovative Approach to the Development of Spatial Data Infrastructures in Europe*, California, ESRI Press.
- Graig, W. J. (1995) *Why we can't share data : Institutional Inertia*, New Brunswick, Centre for Urban Policy Research.
- Grant & Williamson (2003) *State SDI initiatives*, London, Taylor and Francis.
- Gray, B. (1989) *Collaborating - Finding Common Grounds in Multiparty Problems*, Josey-Bass Publisher, San Francisco, CA.
- Grix, J. (2004) *The foundations of Research*, New York, Palgrave Macmillan.
- Groot, R. & McLaughlin, J. (2003) *Geospatial Data Infrastructure*, UK, Oxford Press.
- H.Meredith, P. (1995) *Distributed GIS : If its time is now, Why is it resisted?*, New Brunswick, The State University of NEw JERsey.
- Harvey, F. & Tulloch, D. (2004) *How do local government share and coordinate geographic information : issues in the US*. 10th EC GI and GIS Workshop. Poland.
- Heywood, I. (2003) *Beyond Chorley*, UK, Hobbs the Printers Ltd.
- Kagioglou, M., Cooper, R., Aouad, G., Hinks, J., Sexton, M. & Sheath, D. (1998) *A Generic Guide to the Design and Construction Process Protocol*.

- Kevany, M. (1995) A proposal structure for observing data sharing, New Brunswick, Centre for Urban Policy Research.
- Man, W. H. E. D. (2007) Research and Theory in Advancing Spatial Data Infrastructure Concepts, California, ESRI Press.
- Masser, I (1999) All shapes and sizes : the first generation of national spatial data infrastructure. *International journal of Geographic Information Science*, 13, 67-84.
- Masser, I (2007) Building European Spatial Data Infrastructure, California, ESRI Press.
- Masser, I (2007) GIS World : Creating Spatial Data Infrastructure, California, ESRI Press.
- Masser, I. (1998) Government and Geographic Information, UK, Taylor and Francis.
- Masser, I. (1998) The first generation of national geographic information strategies. 3rd GSDI Conferences.
- Masser, I. (2007) Building European Spatial Data Infrastructure, California, ESRI Press.
- McCullagh, M. (2006) UK SDI Why? What? When? , AGI.
- McDougall, K. (2007) A mixed-method approach for evaluating spatial data sharing partnership for spatial data infrastructure development, California, ESRI Press.
- McDougall, K., Rajabifard, A. & Williaimson, I. (2005) Understanding the motivations and capacity for SDI development from the local level. 8th International Conference on Global Spatial Data Infrastructure. Cairo, Egypt.
- McDougall, K., Rajabifard, A. & Williaimson, I. (2007) A Mixed-Method for Evaluating Spatial Data Sharing, California, ESRI Press.
- Montalvo, W. (2003) Insearch of rigorous models for policy oriented research : a behavioral approach to spatial data sharing. *URISA*, 15, 19-28.
- Nedovic-Budic, Z. & Godschalk, D. (1994) Implementation and management effectiveness in adoption of GIS technology in local government.
- Nedovic-Budic, Z. & Pinto, J. (1999) Understanding Inter-Organisational GIS activities : A Conceptual Framework. *URISA*, 11, 53-64.
- Nedovic-Budic, Z. & Pinto, J. (2001) Organisational GIS interoperability : lesson from the US *International journal of Geographic Applied Earth Observation and Geoinformation*, 3, 290-298.
- Nedovic-Pinto (2004) GIS Database Development and Exchange : Interaction Mechanism and Motivation. *URISA*.
- Nerbert, D. (2004) SDI Cookbook.
- Obermeyer, J. N. & Pinto, J. K. (1994) Managing Geographic Information Systems, New York, Guilford Press.
- Obermeyer, N. & Pinto, J. (1994) Managing Geographic Information Systems, New York, The Guildford Press.
- Oliver (1990) Determinants of inter-organisational relationship, integration and future directions. *Academy of Management Review*, 5, 24-65.
- Omarn, E.-S. E., Bregt, A. & Cromptvoets, J. (2007) Spatail Data Sharing : A Cross-Cultural Conceptual Model, California, ESRI Press.
- Omran E.-S. E., Bregt, A. & Cromptvoets, J. (2006) Spatial Data Sharing : A cross-cultural conceptual model. GSDI conference. Santiago, Chile.
- Onsrud, H. J. & Gracia, M. (2003) Introduction to the special issues on access participatory approaches in using geographic information. *URISA*, 15, 5-7.

- Onsrud, H. J. & Rushston, G. (1995) Sharing Geographic Information System, New Brunswick, Centre for Urban Policy Research.
- Perry (1998) Process of a case study methodology for postgraduate research in marketing. *European Journal of Marketing*, 32, 785-802.
- Pinto, J. K & Onsrud, H. J. (1995) Sharing Geographic Information across organisational boundaries : a research framework, New Brunswick, Centre for Urban Policy Research.
- Pinto, K. J. & Azad, B. (1994) The role of Organisational politics in GIS Implementation. *URISA*, 35-61.
- Prefontaine, L., Ricard, L., Sicotte, H., Turcotte, D. & Dawes, S. (2000) New models of Collaboration for Public Service Delivery. *Worldwide Trend*.
- Prefontaine, L., Ricard, L., Sicotte, H., Turcotte, D. & Dawes, S. (2003) New Models of collaboration for public service delivery - critical success factors.
- Rajabifard, A., Feeney, F. & Williamson, I. (2003) Spatial Data Infrastructure : Concept, Nature and SDI Hierachy, UK, Taylor & Francis.
- Reeve, D. & Petch, J. (1999) GIS Organisations and People : A socio-technical Approach, UK, Taylor & Francis.
- Sexton (2004) PhD workshop: Axiological purposes, ontological cases and epistemological keys.
- Warmest, M. (2005) Collaboration model for National Spatial Data Infrastructure in Federated Countries Department of Geomatics. Australia, University of Melbourne.
- Wegner, M. & Masser, I. (1996) Brave New GIS Worlds, UK, Taylor & Francis.
- Williamson, I. (2003) SDIs - Setting the scene, London, Taylor & Francis.
- Williamson, I., Rajabifard, A. & Binnis, A. (2007) The Role of Spatial Data Infrstructures in Establishing an Enabling Platform for Decision Making in Australia, California, ESRI Press.
- Yin, K. R. (1994) Case Study Research : Design and Methods, UK, Sage Publication Inc.
- Yin, K. R. (2003) Case Study Research : Design and Methods, UK, Sage Publication Inc.

# Towards a Research Design to Investigate Assessment Practices in Built Environment Undergraduate Education

Lloyd Scott and Christopher Fortune, *School of Built Environment, University of Salford, Salford*

## Abstract

'To understand is hard, once one understands, action is easy'  
(Sun Yat Sen, 1866-1952)

Research in Built Environment has begun to emerge as a distinct field. Within that context the investigation and exploration of assessment practices has received very little attention, particularly evident in the area of formative assessment. This type of assessment and the use of effective feedback mechanisms has been an area of interest for this researcher. The aim is to improve the quality of student learning in Built Environment undergraduate programmes through the development of a theoretical model of formative assessment.

The paper discusses the philosophical paradigm which forms the basis for the main research and reflects on the early discourse around the qualitative/quantitative considerations of the researcher. It addresses the philosophical issues surmounted in the choice of research design. The application of a grounded theory approach and more particularly a constructivist stance to the research is explored and rationalised. The key characteristics of this approach are assessed and the advances in conducting and evaluating this design recognised.

The results of the initial research, which give the views and preferences of senior academics, are presented and help inform the next stages of this work in progress. The ongoing work anticipates developing a grounded model for the formative assessment of Built Environment undergraduates for the enhancement of student learning.

Keywords: built environment, formative assessment, grounded theory.

## Introduction

*"What the student does is actually more important in determining what is learned than what the teacher does"* (Shuell, 1986)

This paper focuses on some of the wide-ranging changes that have taken place in assessment practices employed in Higher Education. It looks at assessment in general and then proceeds by setting out its context within the Built Environment, defining it as a theoretical entity. The research design for the research enquiry is considered and explored and the chosen methodology explained and defended. The scope of the research project is offered along with the results of the initial phase. Some early considerations are presented and the next stage of the process identified, and the question of how the initial research might impact on the research project is considered.

## The changing landscape of Higher Education

The drivers of change in HE are numerous and the pressures for that change are occurring globally. Higher education in Ireland has not been ignored on this front. Changes have been brought about in quality assurance arrangements; a National Framework of Qualifications has been introduced and at institutional level many HE establishments have a stated objective of enhancing the student learning experience. An example of this, the Dublin Institute of Technology, a multi-disciplinary provider of Higher Education in Ireland, has put in place a strategic imperative to develop a multi-level learner-centered learning environment through the roll out of a modular structure. A new learning environment was supported by the National Qualifications Authority of Ireland (NQAI) requirement that all awards should be defined in terms of learning outcomes, the achievement of which would be confirmed through the use of appropriate assessment strategies. The traditional approaches to assessment in HE typically place heavy reliance on tacit understanding of standards and in the current environment of rapidly changing contexts this can be a point of strain. Examples of the rapidly changing contexts which have encouraged practitioners to look at more innovative approaches to



assessment include massification, new kinds of teaching and learning, computer-aided assessment, new approaches to intended learning outcomes and declining resources.

### **The impact of assessment on student learning**

Assessment practices in HE have been undergoing wide-ranging changes over recent years and this has been particularly evident in a number of disciplines. These changes are in response to stimuli including a move towards greater accountability, new developments in the use of learning technology and concerns about what graduates need to know, to understand and to be able to do following graduation. The discipline of the Built Environment has been receiving attention in this regard and the validity and effectiveness of traditional modes of assessment have begun to receive consideration.

Assessment in HE is a very complex business, and since assessment is something that is experienced by almost all involved in HE it is important that an assessment system is recognisable and understood by all. There are many reasons to assess students and Brown et al (1996) discuss ten such reasons. Five of these refer to traditional summative assessment and the need for evidence and the classification of learning. The other five focus on formative assessment through guidance for improvement; providing opportunity for students to rectify mistakes to diagnose faults; motivation; providing variety in assessment method and providing direction to the learning process. This might imply that equal importance is placed on both formative and summative, but this is not the case. An investigation of the assessment practices in undergraduate programmes in Built Environment indicates that while the 'tide is starting to turn' there is still an over reliance on the traditional summative examination at the end of a module or unit of learning.

The seminal research material on formative assessment and the use of feedback mechanisms indicates that these methodologies have begun to be recognised as a driving force for enhancing student learning. This has yet to have a complete impact at programme or module level in many undergraduate BE programmes. Research literature informs us that assessment is most effective when it is closely aligned to the learning outcomes. Cross (1996) refers to assessment and feedback as providing one of three conditions for learner success. It is generally acknowledged that a student's approach to learning and the quality of learning achieved will be influenced by the way in which this learning is to be assessed (e.g. Gibbs, 1999; Entwistle and Ramsden, 1983). In addition, adopting a holistic approach to curriculum design that seeks to constructively align assessments with the learning outcomes, and teaching and learning methods that create a seamlessly inter-related curriculum (Biggs, 1999) are important if a diversity of desired learning outcomes is to be achieved (e.g. Gibbs, 1999). Boud (1995) also identifies a need to move from seeing particular assessments in isolation towards recognising them as linked to the other kinds of assessment used, the proximity, frequency and also the context of their usage. Furthermore, bunching of similar types of assessment at certain key points, perhaps at the middle and end of programmes, is likely to result in students' adoption of a surface approach and the attainment of a limited number of lower-level learning outcomes (Scouller, 1996). In other words, cross programme strategic planning of appropriate assessments is fundamental if the intention is for students to attain higher-level learning outcomes such as problem solving and critical thinking (Biggs, 1999; Gibbs, 1999). The critical importance of formative assessment (assessment that contributes to the student's learning through the provision of feedback about performance; Yorke, 2003) should not be underestimated by lecturers and is confirmed by the review work of Black and Wiliam (1998).

Assessment for learning, more commonly understood as formative assessment, is defined by Black and Wiliam (1998: 22) as "all those activities undertaken by teachers and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged". In very simple terms, assessment may be defined as such activities that measure student learning. Boud (1990) posited that assessment has two purposes, firstly that of improving the quality of learning where learners engage in activities and are given feedback that will direct them to effectiveness in their learning (commonly referred to as formative feedback). The second concerns that of the accreditation of knowledge or performance, which occurs generally for the award of a degree or diploma (commonly referred to as summative assessment).

Today, students are more focused and they approach assessment with a better understanding of what is involved. Bloxham and Boyd (2007: 19) refer to students as “being cue conscious concentrating on passing an assessment”. We now hear academics speak in terms of formative and summative assessment, however we have a long way to come before assessment and feedback become central to learning and, in turn, to the student experience. With the importance of life-long learning beginning to permeate thorough HE, along with the impact of the National Frameworks of Qualifications in Ireland, a greater, more explicit emphasis and attention is being paid to learning outcomes and competencies. A student-centred learning framework puts the learner at the centre of the learning process, in which assessment plays an important part. It is widely accepted that assessment has a direct impact on students’ learning (Askham, 1997; Black and Wiliam, 1998; Stiggins, 2002). We are all familiar with the term that *assessment drives learning*; this is true in many instances, where the learner looks at what has to be learned in terms of what he or she needs to do to pass the assessment and get a good grade. Research indicates that what students focus on during the course of their studies is hugely influenced by the assessment methods employed to measure the learning experienced (Ramsden, 1992).

Therefore, the importance of taking cognisance of assessment *for* learning and assessment *of* learning has relevance for lecturers in the design of their assessment strategies. Assessment of learning is where assessment for accountability purposes is paramount; its function is to determine a student's level of performance on a specific task or at the conclusion of a unit of teaching and learning. The information gained from this kind of assessment is often used in reporting and is purely of a summative nature. However, assessment *for* learning, on the other hand, acknowledges that assessment should occur as a regular part of teaching and learning and that the information gained from assessment activities can be used to shape the teaching and learning process. It can, most importantly, also be used by the learner to enhance learning and achievement. Gibbs and Simpson (2004) have developed a model that promotes eleven conditions under which assessment supports learning, as outlined in Table 1 below. Seven of the eleven conditions refer to feedback.

**Table 1:** Gibbs and Simpson (2004) promoting 11 conditions under which assessment supports learning

1. Sufficient assessed tasks are provided for students to capture study time
2. These tasks are engaged with by students, orienting them to allocate appropriate amounts of time and effort to the most important aspects of the course
3. Tackling the assessed task engages the students in productive learning activity of an appropriate kind
4. Assessment communicates clear and high expectations
5. Sufficient feedback is provided, both often and in enough detail
6. The feedback focuses on students’ performance, on their learning and on actions under the students’ control, rather than on the students themselves and on their characteristics
7. The feedback is timely in that it is received by students while it still matters to them and in time for them to pay attention to further learning or receive further assistance
8. Feedback is appropriate to the purpose of the assignment and to its criteria for success
9. Feedback is appropriate, in relation to students’ understanding of what they are supposed to be doing
10. Feedback is received and attended to
11. Feedback is acted upon by the student

## **The Built Environment**

While not the main focus of this paper it is necessary to consider and conceptualise the field of Built Environment (BE). Human society has found it necessary to categorise the different forms of knowledge since well back to the times of Aristotle and Plato in an attempt to make the world more intelligible. Those associated with the BE are no different in this regard. It has begun to emerge as a distinct discipline in the more recent past; however in that discourse it has been identified as problematic. Boyd (2007) refers to the general conception of the BE as one of a 'development process' and he argues that it does not exist theoretically. Ratcliffe (2007), on the other hand, proffers that while the BE is both vague and elusive it is a generic phrase of distinction and pertinence and is best portrayed and understood 'as a set of processes' rather than one single entity. This set of processes includes the planning process, design process, construction process, regulatory process, financial process, transportation process and information process. Griffiths (2004) describes it as a range of practice-orientated subjects concerned with the design, development and management of buildings, spaces and places'.

In HE the field of BE has begun to make significant headway as a recognised discipline where schools of Built Environment have been set up and begun to flourish. The UK Research Assessment Exercise sub-panel makes reference to the field as encompassing 'architecture, building science and engineering, construction and landscape urbanism' (HEFCE, 2005). While school and department configuration is often a matter of the culture of a Higher Education institution, reference to BE by the RAE is acknowledgement of the existence of this discipline. In the Irish HE context, while considered very much at a developmental stage, the field of BE has begun to be recognised and embedded as a distinct discipline. Again, schools and faculty have emerged in the organisation structure of Higher Education institutions across the country.

For the purpose of this research the BE refers to the disciplines of architecture, architectural technology, construction management and construction economics. These disciplines will be the focus of the research as they are the most representative group in terms of BE programmes offered in HE on the island of Ireland. In all the main providers of BE education at undergraduate level, the above areas are offered.

## **Rationale for research design**

Human beings have always shown an interest and concern to come to terms with their environment and to try to make sense and understand the nature of the phenomena to their senses (Cohen et al, 2001). At the commencement of any research project many question occupy the thought of the researcher. What does this journey entail? Where to start? What philosophical stance should one take? What research methods should be employed to effectively achieve the goal(s) of the research? All research needs to be subjected to careful methodological assessment and reflection while theory provides the discourse and a vocabulary to describe what we think. In this regard, the principal aim of the research is to help to improve the quality of student learning in Built Environment undergraduate education. The central research question therefore can be summarised as:

Are assessment practices currently in use in BE education maximising their potential to improve the quality of students' learning?

In attempting to address the aim of the research several research questions are posed:

- How are academics in BE education currently assessing learning?
- To what extent do academics align their assessment practices to educational theory?
- Are the institutional procedures around assessment in conflict with the embedding of a student-focused assessment strategy?
- What are students' experiences and perceptions of assessment?

- What are students' experiences of formative assessment and feedback?
- To what extent do the existing assessment methods encourage a deep approach to learning?
- Do students get an opportunity to reflect on their learning?
- What model can be developed that will enhance the experiences of students with respect to assessment?
- How will the improvements brought about by this new model be measured?

A research framework gives the theoretical background to a research project and most researchers take time to 'struggle' and come to terms with the theoretical aspects of the task. Saunders et al Research Onion model (2003: 87) provides an appropriate form within which to frame this research inquiry. Traditional research design strategies usually rely on a literature review leading on to the formation of a hypothesis which can be put to test by experimentation in the real world. The use of ethnographical and case study approaches can, however, limit the researcher. Grounded Theory (GT), on the other hand, investigates perceived actualities in the real world and analyses that data with no preconceived hypothesis (Glaser and Strauss, 1967). Creswell (2008) offers indicates types of GT, the systematic procedure associated with Strauss and Corbin (1998), the emergent design aligned with the Glaser (1992) and the constructivist approach espoused by Charmaz (2006). The constructivist GT approach which is positioned between the more positivistic stances of Glaser & Strauss and Corbin and the post-modern researchers Lyotard, Foucault and Derrida, and in the camp of Charmaz and Bryant, who question the importance of method, is favoured. The focus on gaining an understanding of the meaning the participants have is an important factor in this research and hence a constructivist bias. Their views, values, assumptions and ideologies with respect to assessment in the BE education are what are sought. The research process considered and developed to address the methodological requirements of a GT approach (Bryant, 2002). The basic GT guidelines are adopted in line with twenty-first century methodological approaches and assumptions. The analysis of data from each of the interview phases, along with data gathered from the survey of academics will influence the emergent theoretical model.

### The first phase of the research process

In the first phase of the research semi-structured interviews were conducted with five senior academics in management positions between September and November 2008 from Schools in the University/Institutes of Technology sector around the island of Ireland. The interviews lasted up to one hour and were taped and transcribed with all interviewees acceptance to be recorded. In a GT approach analysis involves the assignment of concepts and themes to the data gathered, a process recognised as coding. This process was adopted in the case of the data from the interviews. From the analysis the emerging themes and concepts are identified in Table 2 below from this first phase of the research.

**Table 2:** Concepts and codes arising from the first phase of the research enquiry

Concepts	Open Codes
Purpose of assessment	Examination, coursework, regulations, assessment criteria, policies and procedures, summative assessment, formative assessment, holistic assessment, compliance,
Learning and teaching	Teaching methods, improve student learning, innovative practice, scaffolding, reflective learners, modularisation, semesterisation, constructive alignment, student centred learning, independent learning, over assessment, modules, active learning

Academic	Changing practice, learning outcomes approach, traditionalists, coursework, staff development, innovation, course board,
Summative assessment	Examination, coursework, portfolio, measurement, variety, practical tests, peer assessment
Formative assessment	Importance of formative assessment, student involvement, peer assessment, feedback, continuous assessment, portfolio, flexibility,

One emerging concept that is very much identifiable among the Heads interviewed is the difference in philosophical position with respect to assessment and how they view the assessment of student learning. The analysis of the data reflects differing positions as evidenced by the quotes below:

*'The academic staff of the School working with the students are not looking at just the final product as presented but are looking at the process by which the final product was arrived at'* Interviewee A

*'From a management perspective .... I see it being engaged a lot with the compliance with National Framework of Qualifications and adopting changes in relation to learning outcomes process'* Interviewee B

*'This is because we have always proportioned our assessment into end of year exams and coursework'* Interviewee C

Tradition and academic discipline influence the attitude towards the approach to assessment, while the type of educational organisation too has a distinct impact.

The importance of assessment in the educational process was alluded to by all and that formative assessment has an important part to play in this. However the mechanism on how this is achieved differed between each Head. There is a disparity of understanding in the purposes of assessment, particularly as we move towards a more student-centred learning environment. This is evidence by the approach taken in the different institutions with respect to the design of the assessment strategies at module and programme level. There are elements of re-formulating position based on the learning outcomes paradigm in which we find ourselves. For example the interviewee B stated:

*'what we haven't done is link assessment methodologies to module learning outcomes'*

This further emphasises the traditional approach adopted in many programmes and the reliance on the measurement assessment strategy as opposed to a more holistic strategy.

There is evidence of student engagement in active learning tasks as referred to by interviewees 3, 4, and 5, however those tasks are not linked to the overall assessment strategy. Students are required to take a summative exam at the end of a module where they may have demonstrated the achievement of the learning outcomes during the active learning tasks. A clear example of 'over assessment' and a reliance on the traditional summative examination. This position reflect the polarised position across academic institutions in their advancement to the more 'constructively aligned model' advocated by Biggs (1999). This is a common position not just in the BE but across many other disciplines as academic engage in reflecting on and introducing a learning outcomes based approach. One interviewee (4) indicated that academic lecturing staff are unaware that they are 'empowered' to make the appropriate changes to effect learning and hence the more traditional approaches are the preserve. There is still an over reliance on the 'formal summative assessments' or controlled examination as a means of verifying student attainment.

There is clear tension between the summative assessment and the formative assessment processes and using this knowledge/ information to help teaching and learning. Again, the diverse position of each school along the continuum is very much in evidence. In some

instances there has been full engagement in the alignment of programme and module learning outcomes while other schools have only just begun to grapple with this. This one would feel has a direct relationship with their approach and configuration of the assessment strategies employed. This is allied to a complete agreement of the need to strengthen the processes of assessment and in particular the formative assessment elements. The down side is there is no real sense or vision of how this might be achieved. The notion of developing reflective practice through assessment and its contribution to enhancement of student learning and motivation is referred to.

Student involvement in assessment where the academics can benefit from the use of peer assessment on various levels was identified as problematic. The analysis suggests that it happens in a very limited amount of cases. Interviewee 5 indicated that students '*do not perhaps participate as much as they should and that there should be more opportunities to engage the learner more*'. The often laborious process of marking student work can be potentially reduced if some of the assessment is carried out by the students. More fundamentally it can be used to open meaningful dialogue about the work and enhance feedback opportunities. Time constraints and the difficulties associated with peer assessment are cited as the issues associated with engaging students in the assessment. The risks of involving students in their own or colleagues' assessment should not be underestimated. There is intense pressure on the higher education sector to maintain standards. Any change to assessment practice must be able to withstand scrutiny and above all be rigorous and transparent (Race, 2001). There are fears that putting assessment in the hands of the students will make the assessment less reliable. To ensure consistency, measures can be built in, including multiple assessment of the same piece of work by a number of students. Clear definition of marking criteria is another essential element of successful peer assessment. Criteria may be developed with students, but if this is not possible, at the very least they must be made clear prior to students attempting the exercise.

Another emergent theme was the need for inter and intra collegial discussion/ discourse opportunities to discuss not only assessment practices in the Built Environment but also other pertinent pedagogic matters. Ways should be explored of how we might share best practice and how this might begin to effect change in the discipline. This emerged where interviewees made comment on the need to for staff development and training.

### **Concluding thoughts**

This paper has provided a summary overview of the author's research to date with the scrutiny of senior academics regarding their views and experiences in the context of assessment practices in undergraduate Built Environment education. As this is work-in-progress, the paper focused on the methodology employed and a number of key issues emerging from the segment of data analysed thus far. There is a strong history of assessment in the programmes offered in Built Environment undergraduate programmes, particularly the more formal summative assessment. One of the questions to be addressed in the next phase of the process is the extent to which academics are engaging with the most up to date and effective assessment process that will enhance student learning. Interviews with the programme managers and a survey of the built environment academic community will endeavour to address this. The analysis of the results of the initial research set out the views and preferences of senior academics and will help inform the next stages of this work in progress. The ongoing work anticipates developing a grounded model for the formative assessment of Built Environment undergraduates for the enhancement of student learning.

### **References**

- Askham, P. (1997) An instrumental response to the instrumental student: Assessment for learning, *Studies in Educational Evaluation* 23(4), 299-317.
- Biggs, J. (1999) *Teaching for Quality Learning in University*, Open University Press.
- Black, P. and William, D. (1998) Assessment and classroom learning in *Assessment in Education*, 5:1, pp. 7-74.

- Bloxham, S., Boyd, P. (2007) *Developing Effective Assessment in Higher Education*, Open University Press
- Boud, D. (1990) Assessment and the promotion of academic values, *Studies in Higher Education*, 15:1, pp101-111.
- Boud, D. (1995) Assessment and Learning: contradictory or complimentary, in Knight, P. (Ed), *Assessment for Learning in Higher Education*, London: Kogan Page.
- Boyd, D. (2007) Searching for a unified theory of property and construction. In Kosela, L and Roberts, P. (Eds), 'Towards the Foundation of Theory for the Built Environment Symposium' 18-19 June 2007, University of Salford, Research Institute for the Built Environment
- Brown S., Race P., & Smith B., (1996) *500 Tips on Assessment* London: Kogan Page
- Bryant, A. (2002) Re-grounding grounded theory *Journal of Information Technology Theory and Application*, 4(1) pp25-42
- Charmaz, K. (2006) *Constructing Grounded Theory A Practical Guide Through Qualitative Analysis* Los Angeles Sage
- Cohen, L., Manion, L., and Morrison, K. (2001) *Research Methods in Education* (5<sup>th</sup> Ed.) Routledge / Falmer : London & New York
- Creswell, J. W (2008) *Educational Research: 3<sup>rd</sup> Edition* New Jersey: Pearson Prentice Hall.
- Cross, K.P. (1996) Improving teaching and learning through classroom assessment and classroom research, in Gibbs, G. (Ed), *Improving student learning using research to improve student learning*, K., Stanley, N., Davison, N. (Eds), *Teaching in the Disciplines/ Learning in Context*, pp108-113.
- Dainty, A. (2008) *Methodological pluralism in construction management research* in Knight, A. and Ruddock, L. *Advanced Research Methods in the Built Environment* United Kingdom Wiley-Blackwell
- Entwistle, N.J., Ramsden, P. (1983) *Understanding Student Learning*, London: room Helm.
- Gibbs, G. (1999) Using assessment strategically to change the way students learn in Brown, S., Glasner, A. (Ed) *Assessment Matters in Higher Education: choosing and using diverse approaches*, Buckingham: Open University Higher, pp 41-53.
- Gibbs, G. and Simpson, C. (2004) Conditions under which assessment supports student learning. *Learning and Teaching in Higher Education 1*, pp3-31
- Glaser, B. G. (1992). *Basics of grounded theory analysis: Emergence vs. forcing*. Mill Valley, CA: Sociology Press
- Glaser, B. and Strauss, A. (1967) *The Discovery of Grounded Theory* Chicago, Aldine
- Griffiths, R. (2004) knowledge production and the research- teaching nexus: the case of the built environment disciplines *Studies in Higher Education*, 29(6) pp709-726
- HEFCE (2005) *RAE 2008 Consultation on Assessment Panels' Draft Criteria and Working Methods – UOA30, Architecture and the Built Environment* London: higher Education Funding Council of England
- Race, P. (2001) A briefing on self peer and group assessment. LTSN (Learning and Teaching Subject Network) Generic Centre, Assessment Series 9 URL: <http://tinurl.com/tytsswq> (accessed: 31 January 2008)
- Ramsden, P. (1992) *Learning to Teach*, London: Routledge.
- Ratcliffe, J (2007) Built Environment Futures: adopting the foresight principle in formulating and applying a theoretical approach towards the creation of a sustainable built environment. In Kosela, L and Roberts, P. (Eds), 'Towards the Foundation of Theory for the Built Environment Symposium' 18-19 June 2007, University of Salford, Research Institute for the Built Environment
- Sadler, D.R. (1989) Formative assessment and the design of instructional systems, *Instructional Science*, 18, pp145-165.
- Saunders, M., Lewis, P., Thornhill, A. (2003) *Research Methods for business Students*, Prentice Hall

- Scouller, K. (1996) Influence in assessment methods on students' learning approaches, perceptions and preferences: assignment essay versus short answer questions, *Research and Development in Higher Education*, 19:3, pp776-781.
- Shuell, T. (1986) Cognitive conceptions of learning, *Review of Research in Education*, 19, pp 405-450
- Stiggins, S. (2002) Assessment crisis: the absence of assessment for learning *Phi Delta Kappan*, 83:10, pp 758-765
- Strauss, A. and Corbin, J. (1990) *Basic of Qualitative Research: Grounded Theory Procedures and Techniques* Newbury Park CA Sage Publications
- Yorke, M., (2003) Formative assessment in higher education: moves towards theory and the enhancement of pedagogic practice, *Higher Education*, 45, pp 477-501.



# The Hospital as a Physical Space: Accommodating Children's needs

**Elham Sfandyarifard, Patricia Tzortzopoulos and Monty Sutrisna**, *Research Institute for the Built Environment, University of Salford.*

## Abstract

The built environment in healthcare should provide an atmosphere that promotes healing conditions. Despite the general understanding of the impact of the built environment on users, little attention has been given to the children's hospital design from children's perspectives. In this light, the aim of this paper is to investigate factors relevant to supporting the development of child-friendly hospital environments. The adopted research methodology at this stage is literature review. In order to support the creation of child friendly hospital environments, it has been considered important to know the needs of children in a hospital environment. The result shows that different children have different needs and therefore, the design of hospital should ideally be based on diverse requirements of different means of grouping such as age, gender and culture. This paper reports an ongoing PhD research project, which aims to develop strategic guidelines for children's hospital based on children's perspectives.

Keywords: children's needs, children's hospital, Physical environment, wellbeing

## Introduction

Recent years have seen a great amount of attention paid to the built environment and its impact on health care within hospital spaces (e.g. Sundstrom et al., 1996; Lawson, 2001; and Codinhoto et al., 2008). The impact of the built environment on health and recovery from adult's perspective is well known and established (Scher, 1992; Evans and McCoy, 1998).

The Platt report (1959) postulated the differences in children's needs to that of adult's. According to Qvortrup (1995), children are spending most of their time inside home, school or other childhood institutions. There is some research about children and built environments in different spaces such as home (Sibley, 1995) and school (Kraftl, 2006). However, these researches are based on adults' reflections on spatial aspects and built environments of childhood (Birch et al, 2007).

The aim of this research is to develop guidelines for designing a children's hospital supporting health and wellbeing taking into consideration children's preferences. As a part of the ongoing research, this paper intends to explore children's needs in hospital as presented in literature. The question forming this research is:

- What are the physical characteristics of a good children's hospital environment?

This research will be based on a combined research approach, including in-depth literature study and semi-structured interviews within case studies strategy. A theoretical framework will be confirmed from existing literature and will be investigated in different case studies.

This paper is divided into three parts. The first part deals with Children's Hospital Physical environment. It explains what kind of built environment they need to have. The second part presents the research methods used to conduct the research. The third and last session presents the conclusions and further research.

## Children's Hospital Physical Environment

A considerable amount of studies have been performed on the physical elements of healthcare environment and the health outcomes (e.g. Proshansky et al 1976, Sundstrom et al 1996, Lawson 2001). These studies considered isolated elements of the built environment such as lighting, ventilation, colour, and heating as well as the overall design of healthcare buildings. The total environment is experienced by five senses: sight, touch, hearing, taste and smell. These senses help people understand and navigate their environment. All the information that is picked up by a child's sensory receptors will be converted in to a form that the brain can

integrate, interpret, compare and store (NHS Estates, 2003). "Understanding the true dimensions and limitations of the senses enables the designer to create a truly responsive environment that can assist in healing" (NHS Estates, 2003: 45).

Different elements of the built environment related to the five senses of children should be paid attention to during the design of the healthcare space.

### **Initial Literature review**

The effect and influence of children's perceptions on the built environment of the hospital can create a child-friendly environment. However, in order to optimise this effect extra care should be taken in the design procedure of each single element of a children's hospital. Colour, light, noise, artwork, temperature and gardens are the significant parts of the hospitals, which should be designed and employed according to its users.

"The psychological reaction to colour does not preclude the basic biological reaction that stems from human evolution" (Engelbrecht, 2003: 2). According to different studies, the effect of colour on mood, mental clarity and energy levels is inevitable and should be considered as one of the significant design elements (Pile 1997, Camgöz et al. 2003, Engelbrecht, 2003). Such an influence is varied for different types and ages of people. Younger children mostly favour bright colours, whereas adolescents prefer lower toned colours (Engelbrecht, 2003). Gender also plays a part in colour and design preferences among children of different age groups.

Attention to lighting is also seen as central to the creation of therapeutic environments for adult and children (Birtch et al., 2007).

There is a considerable amount of research conducted on lighting and its effect on young patients in hospital (e.g. Moules and Ramsay, 1998; Pölkki et al., 2001). Lighting has been known and used as a powerful tool especially in hospitals. Some examples of lighting therapies are: (a) phototherapy used in the treatment of neonatal jaundice instead of blood transfusions; and (b) Ultra violet light for a dermatological condition, such as psoriasis (NHS Estates 2003).

One of the significant elements in children's hospitals is thematic design used in all spaces occupied by people especially patients (Coad and Coad, 2008). According to the studies performed by Coad and Coad (2008), very strong preferences for designs and textures among the young patients exist. The thematic design is associated with age of patients; younger children are more interested in animals and cartoon characters where as children from ages 6 to 10 years prefer the sea having a beach, fish, seashells, boats and creating an idyllic holiday scene and children from 11 years to adolescents prefer conceptuality in patterns and designs such as abstract designs (Dyer, 1999).

According to Warner (1995), during the Middle Ages, when monasteries were the hospitals of the day, the garden was one of the main elements as a place of healing.

Several studies (Strain and Grossman, 1975; Horsburgh, 1995) maintain the importance of the provision of gardens in hospitals and access to nature by hospitalised children. This has the benefit of reducing the stress of parents during the hospitalisation of their children (Carpman and Grant, 1993).

In healthcare built environments, the use of powerful and effective artwork can convey appropriate messages such as hope, joy, love, dignity, peace, tranquillity, energy, comfort, security, safety, growth and life (Dyer, 1999). It can be applied in different ways such as through paintings, murals, prints, photographs, sculptures, decorative tiles, ceramics, textile hangings and furniture.

Ventilation has been identified in different research studies as a major contributing factor in reducing the occurrence and spread of infection and improving other conditions that otherwise would lead to the spread of contagious diseases (NHS Estates 2003, Birch et al 2007). In order to tackle this problem, an appropriately designed air conditioning system is vital. Ventilation design in general view is essential for both children and adults. However, children are more sensitive about different smells (NHS Estates, 2003).

In addition to the air quality of the hospital, the temperature also plays an important role in the healthcare environment. Despite the importance of temperature in a healing environment, not many studies have been done in recent years. According to the Department of Health 2003, inappropriate temperature, cool or warm, gives a sense of illness to the patients both adults and children.

There is a considerable amount of literature concerning the effect of noise on wellbeing (Ulrich 1991, Redshaw 2004, NHS Estates 2004). According to NHS Estates 2003, unwanted sound or noise can increase heart rate, blood pressure, respiration rate and even blood cholesterol levels. The ambient noise threshold in children is lower than in adults. Redshaw (2004) noted that too much noise in hospital for child outpatients could be problematic.

The studies in noisy surroundings in children's hospitals were initially concerned about babies crying and the children especially ages between 4 to 7 years who complain a lot about them. Another line of research interest relates subjective perceptions of noise and noise (L4A 1999; Derbyshire Children's Hospital) suggests that a separate ward for babies and a separate quiet room, to which a crying child might be taken to be comforted.

### **Research methodology**

The study begins with a review of relevant materials from textbooks, journals, conference papers, and Internet information to capture the background of the hospitals environment. According to Hart (1998:.13) the literature review is "the selection of available documents (both published and unpublished) on the topic, which contain information, ideas, data and evidence written from a particular standpoint to fulfil certain aims or express certain views on the nature of the topic and how it is to be investigated, and the effective evaluation of these documents in relation to the research being proposed". Moreover Blaxter et al. (2006) defined the purpose of the literature review as to address the research and to create the context and insights into previous work (Ridley, 2008).

Existing studies were reviewed, and their limitations were identified. By integrating insights from previous studies, a set of characteristics, which are believed to be, common in the design of children's hospital were identified. Finding a gap in knowledge, developing a research framework and establishing future research program were the main purpose of the literature review.

Following this, semi-structured interviews will be conducted within the context of case study strategy. Semi structured interviews have been defined as the sole method used in flexible designs or in combination with others. In this type of interviews, the interviewers have freedom in the sequencing of questions and attention given to different topics through their exact used wording and time(Robson, 2002:.278), whilst case study is defined by Yin as "an empirical inquiry that investigates a contemporary phenomenon within its real-life context; when the boundaries between phenomenon and context are not clearly evident; and in which multiple sources of evidence are used" (Yin, 1994: 23).

This has been considered suitable to satisfy the aim and objectives of this research, i.e. to identify what are the physical characteristics of a good children's hospital environment.

### **Recommendations from the literature review**

The findings of this study have a number of important implications for future practice:

- The physical environment of hospitals should be bright, colourful, tidy, spacious, welcoming, comfortable and quiet
- Controls for the lights and air circulation devices should be accessible for children
- Design of different type of interesting physical environments such as spaces of different sizes, places to hide, natural and man-made things is helpful in the healing process of patient
- Design of the hospital based on consultation with children should be paid careful attention due to their diverse requirements (especially in terms of age and gender).

- Access to each department, ward or public facility should be signposted clearly

## Conclusions and Future Studies

This paper addressed the investigation of children's needs in a healing environment. The physical environment could be designed with sensitive lighting, colour, sound attenuation, ventilation and artworks to improve the child's wellbeing. Therefore, the aim of this paper was to present the finding from the literature on children's needs for a child-friendly hospital. A friendly environment could help to distract a hospitalised child from the reality of their illness. The literature review demonstrated that extra care should be taken in the design procedure of each single element of children's hospital. It is now possible to state that air change ventilation, comfortable parent and family accommodation, creative and colour-coded wayfinding initiatives, clever use of time-appropriate interior lighting, diverting art, varied and age-appropriate play facilities and multiple gardens and outdoor recreation spaces are the characteristics of a child-friendly hospital environment and have a major effect on the healing process.

Another gap is that the existing papers are discussing the general character of children hospital in overall and there is not enough evidence about different medical specialisation. Considerably more work will need to be done to formulate guidelines for designing child-friendly hospitals supporting health and wellbeing taking into consideration the children's preferences and perceptions.

The NHS Plan (2000) argued that it is essential to focus more on the things that really matter to patients to improve the patient's experience of the hospital environment. Douglas and Douglas (2002) indicated that the wellbeing of patients depends on the built healthcare environment. Therefore, it is important to ensure that patient satisfaction is not only prioritised but also achieved. Healthcare environments should reflect patients' perceptions and preferences.

NHS Estates (2003) ought to make an effort at ensuring the provision of friendly and welcoming healthcare environments. In addition to that, the Department of Health (DH, 2003) has suggested that the hospital environment should be prepared for children from the child's point of view.

The next step of this research will be an investigation of children's preferences and perceptions in the context of children hospitals and formulating a theoretical framework, which will be investigated in different case studies.

## References

- Birch, J. Curtis, P. James, A. (2007), 'Sense and sensibilities: in search of the child- friendly hospital', *Built Environment* 33:4 405-416
- Camgöz N, Yener C and Güvenc D. (2003), 'Effects of Hue, Saturation and Brightness: Part 2: attention'. *Color Research and Application* 29(1): 20-28
- Carpman, J. R. and Grant, M. A. (1993), *Design that Cares: Planning Health Facilities for Patients and Visitors* (2nd edition). Chicago: American Hospital Publishing.
- Coad J and Coad. N. (2008), Children and young people's preference of thematic design and colour for their hospital environment', *Journal of Child Health Care*, 12: 1, 33-48,
- Codinhoto, R., Tzortzopoulos, P., Kagioglou, M., Aouad, G., Cooper, R. (2008), *The effect of the built environment on health outcomes*, The Health and Care Infrastructure Research and Innovation Centre.
- Department of Health. (1959), *The Welfare of Children in Hospital –Report of the Platt Committee*, The Stationery Office, London
- Department of Health. (2000), *The NHS Plan*. The Stationery Office, London.
- Department of Health. (2003), *Getting the Right Start: National Service Framework for Children. Standard for Hospital Services*, Department of Health Publications, London
- Dyer A. (1999), *Child and Parent Concept Sessions*. Great Ormond Street Hospital for Children. NHS Trust, London.
- Engelbrecht K. (2003), *The Impact of Colour on Learning*, available at: [www.merchandisemart.com/neocon/NeoConConfPro/W305.pdf](http://www.merchandisemart.com/neocon/NeoConConfPro/W305.pdf), accessed 30.11.04.

- Evans, G. W., McCoy, J. M. (1998), 'When buildings don't work: the role of architecture in human health'. *Journal of Environmental Psychology*, 18, 85-94.
- Hart C. 1998, *Doing a literature review*, Sage, London.
- Horsburgh, C. R. Jr. (1995), 'Healing by design'. *The New England Journal of Medicine*, 333(11): 735-40.
- Kraftl, P. (2006), 'Building an idea: the material construction of an ideal childhood'. *Transactions of the Institute of British Geographers*, 31, 488-504.
- L4A (1990) ,*Building the Best. The New Derbyshire Children's Hospital*, report of a consultation workshop with staff, parents and children,
- Lawson, B. (2001), *The language of space*. Elsevier, Oxford
- Moules, T. and Ramsay, J. (1998) *The Textbook of Children's Nursing*. Stanley Thornes, Cheltenham.
- NHS Estates.2003, *Improving the Patient Experience – Friendly Healthcare Environments for Children and Young People*. The Stationery Office. Norwich
- NHS Estates.2004, HBN (Health Building Note) 23. *Hospital Accommodation for Children and Young People*. The Stationery Office. London
- NHS Trust (1999), *Wayfinding*.The Stationery Office, Leeds.
- Pile J, 1997. *Color in Interior Design*, McGraw-Hill. New York
- Pölkki, T., Vehviläinen-J., ulkunen, K. and Pietilä, A.M. (2001), 'Nonpharmacological methods in relieving children's post-operative pain: a survey on hospital nurses in Finland'. *Journal of Advanced Nursing*, 34(4), 483-492.
- Proshansky, H., Ittelson, W. (1976), *Environmental psychology : people and their physical settings*. Holt, Rinehart and Winston. London, NewYork.
- Qvortrup, J. (1995), 'Childhood and modern society: a paradoxical relationship', in Brannen, J. and O'Brien, M. (eds.) *Childhood and Parenthood: Proceedings of the ISA Committee for Family Research Conference on Children and Families 1994*, Institute of Education, London.
- Redshaw, M. (2004), Design for Health. *The Impact of a New Hospital Environment on Children, Families and Staff*, Faculty of Applied Sciences, University of the West of England, Bristol.
- Ridley, D. (2008), *Literature Review: A step by step guide for Students*, SAGE Publications Ltd.
- Robson,C. (2002), *Literature Review: Real world research*, John Wiley and sons Publications Ltd, USA.
- Scher, P. (1992), *Environmental Design Quality in Health Care*. Arts for Health, Manchester.
- Sibley, D. (1995), 'Families and domestic routines: constructing the boundaries of childhood', in Pile, S.and Thrift, N. (eds.) *Mapping the Subject*. London. Routledge, 123-137.
- Strain, J. J. and Grossman, S. (1975), 'Psychological reactions to medical illness and hospitalization', in Strain, J. J. & Grossman, S. (Eds) *Psychological Care of the Medically Ill: a Primer in Liaison Psychiatry*. NewYork: Ap- pleton-Century-Crofts, 23-36.
- Sundstrom, E., Bell, PA., Busby, PL, Asmus, C. (1996). 'Environmental Psychology 1989-1994.' *Annual Review Psychology* 47:485-512.
- Ulrich, R.S. (1991), 'Effects of health facility interior design on wellness: theory and recent scientific research', *Journal of Health Care Design*, 3, 97-109.
- Warner, S. B., Jr. (1995), 'The human side: A brief history of healing gardens in health care settings'. *Journal of Health Care Design*, 7, 213-219
- Yin, K.R. (1994), *Case study research: design and methods*, Sage, London.

# Heaven, Hell and Complementarity: Opposition and Duality in late-Georgian and Regency England

Christine Simon, *Department of English, University of Chester*

## 1.

Richard Brothers (1757-1824) – a retired naval lieutenant on half pay – believed himself to be the nephew of Jesus Christ and styled himself 'Prince and Prophet of the Hebrews'. It was, he claimed, only his intervention with the Almighty in 1791 which averted the destruction of London in a fit of divine wrath. In 1795, he believed he was tasked by God to lead a mass exodus of the lost tribes of Israel (amongst whom were many Britons) from Britain to Palestine, where he was to rebuild Jerusalem and initiate a messianic rule of peace; to this end, he assumed that King George III would abdicate the throne in his favour. Later that year, however, Brothers was arrested and convicted on a charge of criminal lunacy; for the next eleven years he was confined to Fisher House, a private asylum in Islington (Southey, 1984: 427-423) (Thompson, 1991: 127-129) (*Dictionary of National Biography*, 2009).

Nearly twenty years after Brothers' abortive exodus, a sixty-four-year-old virgin named Joanna Southcott (1750-1814) – a former Devon maidservant – claimed to be pregnant with Shiloh, the son of God. This was the culmination of a tempestuous ministry which had started with millenarian visions in 1792 but which did not gain widespread attention till the turn of the century. In Brothers' absence, many of his former disciples turned their allegiance to Southcott, some of them setting her up in London in the early years of the nineteenth century. Upon her many devotees Southcott imposed the dictates of her heavenly 'voice', which often caused her to erupt in outbursts of anger on behalf of the Almighty (Southey, 1984: 433-446), (Thompson, 1991: 420-428) (*Dictionary of National Biography*, 2009).

Both Brothers and Southcott attracted thousands of supporters from amongst the ordinary people. On one level, their ministries can be seen as nothing more than straightforward instances of self-delusion or psychosis; the madcap fringe of a society in crisis; the expression of apocalyptic hope which attracts the desperate in difficult times. This was a period of stark oppositions and upheavals both national and international. The French Revolution had polarized the country, some seeing it as a harbinger of radical reform in Britain, others as a calamitous threat to the natural and God-ordained hierarchies of monarchy, aristocracy and established religion. But the free Britain vaunted by the Revolution's opponents was nonetheless a country divided; not least by the distribution of wealth and the disenfranchisement of the majority of the population: all women and those men who did not have the requisite property qualifications. ('It appears,' declared the London Corresponding Society in its 'Address to the Nation' of 24 May, 1792, 'that 257 supposed Representatives of the People, making a Majority of the House of Commons, are returned by a Number of voters not exceeding the thousandth part of the Nation.' (Thale, 2008: 13)) Repressive laws accompanied by draconian punishments controlled the expression of anything considered by the Government to be seditious. The Napoleonic Wars were to drag on until 1815, using up the country's resources and producing a weariness in all levels of society. Unruly times breed prophets and cranks; Brothers and Southcott were no exception.

However, if we come at them from a different angle, a different picture emerges. These prophets may have drawn bigger crowds and had a higher profile than many others; but, viewed against the religious and political currents of the age, they were not so unusual. During the 1790s, England sparked with an apocalyptic fervour in which both political and religious ideas were mingled, expressed in both learned and popular modes. In 1789, shortly after the outbreak of the French Revolution, the rationalist Unitarian minister Dr Richard Price ended his 'Discourse on the Love of Our Country', a sermon given to the Society for Commemorating the [1688] Revolution in Great Britain, with the following apocalyptic warning:

Tremble all ye oppressors of the world! Take warning all ye supporters of slavish governments and slavish hierarchies! [...] You cannot now hold the world in darkness. Struggle no longer against increasing light and liberality. Restore to mankind their

rights and consent to the correction of abuses, before they and you are destroyed together. (Quoted in Thomas, 1991: 196)

Brothers' cranky mission reflects this mix. It had a political edge which was by no means unusual when seen against the complex blend of religious dissent and radical politics of the time. His prophecies contain a mixture of biblical idiom and late eighteenth-century millenarianism, much of it vague enough to be immediately applicable in such troubled times. In his pamphlet *Brother's [sic] Prophecy of all the Remarkable and Wonderful Events which will come to pass*, he wrote:

The proud and lofty shall be humbled, even to the dust; but the righteous and poor shall flourish on the ruins of the wicked. [...] The pestilence shall sweep away the Locusts that eat up the harvest of Industry; and the Earthquake shall swallow up the monstrous *Leviathan*, with all his train. In all these things the poor, the honest, the virtuous, and the patriotic, shall rejoice. (Quoted in Thompson, 1991: 128)

On the one hand, this is no more than a nebulous regurgitation of biblical sayings. But at the same time, Brothers' concern for the deserving poor can be seen as a reflection of a growing social awareness during this period. Brothers' choice of vocabulary is interesting and perhaps significant. Is 'Leviathan' here merely a vague biblical metaphor, or does it refer to the social machinery, as in Hobbes? Does 'patriotic' have its normal sense of 'loyal to one's country' or is it imbued with the new French nuance of 'pro-revolutionary'? In the inflammatory atmosphere of the 1790s, these potentially revolutionary concepts were not to go unnoticed. Likewise, Brothers' refusal to sign a document which contained the words 'our Sovereign Lord', referring to George III (he considered them blasphemous) (*Dictionary of National Biography*, 2009) could be interpreted as nothing more than a somewhat immoderate religious loyalty, an absolute allegiance to a heavenly king whose sovereignty ought not to be challenged by an earthly one.

However, Brothers was not the only one to be rocking the boat of monarchy during these years. Tom Paine also had a dim view of kingship; in *Common Sense* (1776) he had written:

In England a king hath little more to do than to make war and give away places; which in plain terms, is to impoverish the nation and set it together by the ears. A pretty business indeed for a man to be allowed eight hundred thousand sterling a year for, and worshipped into the bargain! Of more worth is one honest man to society, and in the sight of God, than all the crowned ruffians that ever lived. (Nelson, 2007: 87)

And in *Rights of Man* (1790) he had sung the praises of the French Revolution and extolled the virtues of republicanism:

Conquest and tyranny transplanted themselves with William the Conqueror from Normandy into England, and the country is yet disfigured with the marks. May then the example of all France contribute to regenerate the freedom which a province of it destroyed! (Paine, 1985: 129)

Although the basis for Paine's political and religious ideas was diametrically opposed to that of Brothers – he was a rationalist and a deist – and although those ideas were more fully reasoned and developed into a system, we may discern in both approaches, and their concern with change and social justice, a consonance or a variation on a theme. Paine too was an object of establishment wrath because of his undue influence on the populace; both *Rights of Man* (the second part of which earned its author the death penalty in Britain) and *The Age of Reason* were considered all the more inflammatory because they were aimed at ordinary people. This was a time already disturbed by civil unrest and the threat of popular revolution. By 1795, Pitt's government was struggling with defeats in the war with France, crop failures and popular discontent. Brothers' high profile and the large press coverage that went with it; his religious fervour with its possible political overtones; the fact that he was visited by

members of the radical London Corresponding Society (Thompson, 1991: 129), all meant that his imprisonment had as much to do with what the authorities saw as his destabilizing effect on the populace as it did with his religious views and his 'madness'. This ambiguity surrounding his status is reflected in the government's confusion over the nature of the charge on which Brothers was imprisoned: originally apprehended for treason, this was subsequently changed to criminal lunacy (Madden, 2008: 273).

Joanna Southcott permits a similar dual interpretation. On the surface she is the quintessential religious fanatic, complete with paroxysms, crazy prophecies and bossy dicta. Like those of Brothers, many of her prophecies are nothing more than a mishmash of biblical phrases, superficially adapted to the present situation. The following is typical:

O England! O England! England! the axe is laid to the tree, and it must and will be cut down; ye know not the days of your visitation [...] The midnight hour is coming for you all, and will burst upon you. I warn you of dangers that now stand before you, for the time is at hand for the fulfilment of all things. (quoted in Thompson, 1991: 422)

But Southcott too exemplifies the oppositional patterns of the period. Her ministry was not as overtly political as that of Brothers; by the time of her greatest influence, in the early years of the nineteenth century, millenarianism had – in the wake of the war with France and the clampdown on British radicalism – to some extent retreated once more to the religious arena. And like the Methodists to whom she once belonged, Southcott laid emphasis on a purely spiritual salvation and reward in the next life. However, she also saw herself as a champion of the poor. 'I hear the cries of the poor, complaining they are starving to death, for want of food,' she wrote to the Reverend Pomeroy, vicar of St Kew in Cornwall in 1800 (*Dictionary of National Biography*, 2009). And she (or her 'voice'), in a mix of biblical idiom and topical comment similar to that of Brothers, cursed those leaders ('shepherds') who neglected the plight of the poor:

My charges will come heavy against them, and my judgements must be great in the land, if they starve the poor in the midst of plenty. [...] What I said of Nineveh, Sodom and Gomorrah, what I said of Tyre and Sidon, what I said concerning the Galileans, are now charges against the shepherds of England. (Quoted in Thompson, 1991: 424)

E. P. Thompson even sees in one of her verses what he calls 'an echo' of Paine's line in *Common Sense* which depicts William the Conqueror as the 'bastard and his armed banditti':

For I'll cut off the bastard race,  
And in their stead the true heirs place  
For to possess the land. (Quoted in Thompson, 1991: 423)

Southcott may be seen as a representative of the marginalized, women and the lower classes alike. Her calling bypassed the traditional requirements of spiritual leadership: university education and ordination (i.e. wealth and social standing) on the one hand, and maleness on the other. To some extent this is consonant with the general dissenting experience, which laid emphasis on a personal relationship with God and on the equality of souls (with all the political undertones which that entailed); but in her insistence that God had chosen not only a member of the lower classes but a woman to be his mouthpiece, she parts company even with most of the dissenting sects. 'I will conquer in woman's form,' her voice told her in June 1804.<sup>1</sup> Whatever the psychological processes by which Southcott became convinced that she was a channel for God's voice, her ministry is of interest from a feminist standpoint as well as a political one. 'If the woman is not ashamed of herself, the Devil cannot shame her,' she wrote (quoted in Southey, 1984: 423), thus showing herself to be, in some degree at least, a sister to the more rational and enlightened Mary Wollstonecraft.

---

<sup>1</sup> Letter to William Sharp, quoted in *Dictionary of National Biography* (2009).



It also needs to be noted that it was not only 'madmen' and 'old, vulgar, and illiterate'<sup>2</sup> women who shared the ideas of Brothers and Southcott. Although their main appeal was to the poor and the downtrodden, those for whom religion offered the only consolation, they both numbered amongst their devotees members of the educated classes. Nathaniel Halhed, M.P. and Sanskrit scholar, was a vocal supporter of Brothers both in and out of Parliament. William Sharp, a radical engraver, exemplifies the transition of allegiance from Brothers to Southcott which took place after the former's incarceration. Sharp was a friend of William Blake (himself a man of unorthodox views and prophetic vision, albeit on a different intellectual and artistic plane) and part of the circle around the radical London publisher Joseph Johnson which included such figures as William Godwin, Tom Paine and Mary Wollstonecraft. In 1795, he had produced an engraving of Brothers, under which he placed the unequivocal caption, 'Richard Brothers, Prince of the Hebrews: Fully believing this to be the Man whom GOD has appointed, I engrave his likeness.' (*Dictionary of National Biography*, 2009). He was a member of the group of men responsible for Southcott's establishment in London; this group of seven, of whom six were followers of Brothers and three were Anglican clergymen, had travelled to Exeter in order to sound her out; they returned convinced. The well-to-do Jane Townley opened her home to Southcott in London and offered her maidservant to act as Joanna's secretary.

## 2.

We turn now to another aspect of early nineteenth-century culture. In 1803, when Southcott had just begun to establish herself in London and to gain popularity, Thomas Young, a lapsed Quaker and gentleman doctor, gave a lecture to the Royal Society ('On the Theory of Light and Colours') which reignited the debate surrounding the nature of light. Scientific opinion at that time was divided: was light a series of particles (as Newton had implied and most English natural philosophers believed) or a wave (as Huygens had posited)? In his 1803 experiment, Young placed slips of card in a beam of light and observed the diffraction patterns caused by the splitting of the light beam. The same results were obtained by a slightly different experiment, for which Young is more famous: the double-slit experiment, in which a beam of light is passed through two tiny slits and the resulting patterns observed on a wall or screen. These patterns – alternating bands of light and dark – could only be explained if the beams of light were behaving as waves: the light bands caused by constructive interference, the dark ones by destructive interference. (The precise date of the double-slit experiment is uncertain, but it must have been performed after the 1803 lecture and before the publication of Young's *A Course of Lectures on Natural Philosophy* in 1807, in which an account of it is given (Robinson, 2002: 144). The two experiments are, in fact, equivalent (Scheider, 1986)). Both these experiments provided sound evidence for the undulatory theory of light; but Young's conclusions were criticized by Henry Brougham, fellow member of the Royal Society and radical lawyer, with such vehemence and acrimony that Young abandoned the study of light altogether. As the nineteenth century progressed, however, more and more evidence accrued for the wave theory, which was eventually adopted. The work of James Clerk Maxwell (1831-1879) in the 1860s was conclusive. Maxwell unified the theories of electricity and magnetism and showed that light is a form of electromagnetic radiation. 'Light,' he wrote, 'consists in the transverse undulations of the same medium which is the cause of electric and magnetic phenomena.'<sup>3</sup>

But that was not the end of the story. In the twentieth century came the insight of quantum theory that light is not *either* a wave or a particle, but *both*. Young's double-slit experiment was re-enacted in another: the electron double-slit experiment. This was originally a thought experiment devised by Richard Feynman, but was later performed, giving the results Feynman had predicted. In this experiment, electrons – rather than a beam of light – are fired through the two slits. Electrons would be expected to behave as particles and accumulate in two heaps opposite each of the slits. In fact, something very strange happens: although the electrons can be detected as particles as they pass through the slits, by the time they appear at the screen they are behaving as waves, and form an interference pattern similar to that obtained in Young's double-slit experiment with light. This wave/particle duality, or complementarity,

---

<sup>2</sup> Southey's descriptions of Brothers and Southcott respectively (Southey, 1984: 428, 433).

<sup>3</sup> Quoted in 'James Clerk Maxwell', *Encyclopaedia Britannica*, CD-ROM, 2008.

extends not just to light but to all reality. Ask wave questions and you get wave answers; ask particle questions and you get particle answers. Like one of those black and white pictures which reveal an old hag or a beautiful young woman depending on how you look at it, reality is characterized by an inherent ambiguity (Tudor Jones and Wall, 2009: 89-96).

Also at the heart of the quantum world is uncertainty. Although in the double-slit experiment the final pattern is always the same, it is impossible to predict the behaviour of any single electron. The individual motions of the particles are unpredictable. It is possible to predict the behaviour of large numbers of particles (what may be called the 'statistical model') but not that of individual ones (the 'dynamic model'). Heisenberg's uncertainty principle states that it is impossible to know simultaneously the position and the momentum of a particle (an electron, for example). This is not to do with any inadequacy of equipment or of experimental methods, but is an inherent property of nature (Gribbin, 1991:119-120).

It would appear also that, since the electrons are behaving as waves (producing interference patterns) and since the same results are obtained if only one electron is released at a time, each electron somehow passes through both slits (Gribbin, 1991: 170-171). However, if the experiment is set up in such a way that individual electrons are 'observed' (detected) as they pass through the slits, it is found that each electron passes through either one slit or the other, and the interference pattern disappears. What appears on the screen in this case is the pattern of distribution which would be obtained if, for example, bullets were being fired. The act of observation changes the outcome, since it introduces energy into the experiment which changes the state of the object (Tudor Jones and Wall, 2009: 96) 'The observer interacts with the system to such an extent that the system cannot be thought of as having independent existence' (Gribbin, 1991: 160).

Although this uncertainty is insignificant in the behaviour of objects on the everyday level, since the uncertainty involved is minute, it nevertheless has radical implications for our understanding of reality. Because the macro world consists of atoms whose component particles behave according to quantum rules, all material reality is in some sense underpinned by the strange concepts of quantum theory (Gribbin, 1991: 146-147). 'Nothing is real until we look at it, and it ceases to be real as soon as we stop looking,' is how one physicist has put it (Gribbin, 1991:173).

### 3.

So we have come, via cranky turn-of-the-century prophets and the debate on the nature of light, to complementarity and the bizarre world of quantum theory. What does this have to do with my novel-in-progress *Chimera*?

First and most simply, the setting. Part of the novel is set in the age of Brothers and Young, itself an age, as we have seen, of oppositions. Young's 1803 lecture at the Royal Society is attended by one of the historical protagonists. Part of the novel is set in the present day, and thus the two periods are juxtaposed and contrasted.

Secondly, one of the main themes of *Chimera* is that of opposition and duality. The two present-day protagonists, Julia Dalton and Peter Marchmont, are both researching the same historical characters, albeit with different motives and insights. The main historical protagonist, Richard Turnbull, is a shadowy and anomalous character who seems to be neither one thing nor another. A radical and an atheist, a supporter of the French Revolution, as so many British radicals were before the Terror, yet in 1810 he volunteers to work as a spy-catcher for the British government, bringing to justice his friend Henri de Saint-Gilles, who claims he has been falsely accused. The plot of *Chimera* hinges on a problem: was Turnbull a French spy, an English spy, or a double agent? Was he a loyal friend or a traitor?

Finally, I have used the 'both/and' paradigm of the theory of light as a patterning device in the novel. As a metaphor it has some interesting possibilities. As we have seen, phenomena such as the ministries of Richard Brothers and Joanna Southcott can be seen in two opposite ways, both equally accurate. Similarly, Julia Dalton is faced with oppositional readings of the characters she is researching. *Chimera* asks questions concerning the nature of truth and the ambiguities of reality. Is truth singular or multiple? Are the binary oppositions of black and white, good and evil, fact and fiction, the only – or the most beneficial – ways of describing reality? A profound truth, said Niels Bohr, is one whose opposite is also true (Bynum, 2006). Julia Dalton's twenty-first-century insight may be that Richard Turnbull, like light, was not one

thing or another, but both, depending on which questions she asks of the historical record and which questions history asked of Turnbull.

## References

- Bynum, W. F and Porter, R. (eds), *The Oxford Dictionary of Scientific Quotations*, Oxford University Press, Oxford, 2006
- Dictionary of National Biography* (2009), available at <http://www.oxforddnb.com/>
- Encyclopaedia Britannica*, CD-ROM, 2008
- Gribbin, J. (1991), *In Search of Schrödinger's Cat*, Black Swan, London
- <http://www.sciencedirect.com/science/journal/01916599>
- Madden, D (2008), 'The religious politics of prophecy: Or, Richard Brothers's *Revealed Knowledge* confuted', *History of European Ideas*, 34, (2), pp. 270-284. Available at:
- Nelson, C (2007), *Thomas Paine: Enlightenment, Revolution, and the Birth of Modern Nations*, Profile Books, London
- Paine, T (1985), *Rights of Man*, Penguin, London
- Price, R, (1991) *Political Writings*, Cambridge University Press, Cambridge
- Robinson, A. (2006), *The Last Man Who Knew Everything: Thomas Young, The Anonymous Polymath Who Proved Newton Wrong, Explained How We See, Cured the Sick, and Deciphered the Rosetta Stone, Among Other Feats of Genius*, Oneworld, Oxford
- Scheider, W. (1986), 'Do the "Double Slit" Experiment the Way it Was Originally Done', available at <<http://www.cavendishscience.org/phys/tyoung/tyoung.htm>
- Southey, S. (1984), *Letters from England*, Alan Sutton, Gloucester
- Thale, M. (ed.) (2009), *Selections from the Papers of the London Corresponding Society 1792-1799*, Cambridge University Press, Cambridge
- Thompson, E. P, (1991), *The Making of the English Working Class*, Penguin, London
- Tudor Jones, G. and Wall, A. (2009) in A, Wall, *Myth, Metaphor and Science*, Chester Academic Press, Chester, pp. 89-96.

# **An ethnographic study of the culture in a Diagnostic Imaging department (DID) – some initial findings**

**Ruth Strudwick**, *School of Nursing, University of Salford.*

## **Abstract**

This is my doctoral thesis for my Professional Doctorate, which is ongoing. I have completed my data collection and have some initial findings to present. There has been very little written about radiographers and how they work and interact. This study will explore the culture in a DID, looking at how radiographers work and the issues within their working environment. This study will contribute to the sparse evidence base by providing valuable insights into how radiographers work in the pressurised environment of the NHS. The results of this study may prove beneficial for prospective diagnostic radiographers (DRs) and other health and social care professionals. So far the study has involved observation for a four month period in a DID in the East of England. Observation of staff in the DID was carried out to identify current issues for DRs. After a period of observation the researcher conducted semi-structured interviews with key informants from the DID. The interviews explored issues identified from the observation in greater depth.

Documents used in the DID were also analysed. This paper will outline my study so far and present some of the initial findings.

## **Introduction**

Much of the research that has been undertaken in radiography is quantitative research looking at radiographic techniques and imaging methods and modalities. Ng and White (2005) encourage the use of qualitative research to look at the interactions that occur within a DID. They state that qualitative research is needed in radiography to “provide insight into certain topics of which little is known” (p217), for example, looking at the perceptions and experiences of DRs. Adams and Smith (2003) support this idea saying that “there is considerable potential for the sustained use of qualitative methodologies in radiography research to more clearly define what radiographers do and how they do it” (p194).

## **Methodology**

The reason for my choice of a qualitative methodology is that qualitative research inquires into the meaning individuals or groups ascribe to a social or human problem; it allows for the exploration of people’s thoughts, feelings and ideas (Creswell, 2007). The purpose of my research is to investigate the culture in the DID amongst DRs and in order to do this I need to see the culture from the perspective of those who are a part of it, namely the DRs working in that DID (Crotty, 2005). Quantitative research does not provide meanings, it provides numerical data and hard facts (Bowling, 2002). Quantitative techniques can be used if the subject is known about, simple and unambiguous and able to be measured in a valid and reliable way (Bowling, 2002). Qualitative methods provide further insight and rich data about the complex issue of culture (Bowling, 2002).

## **Ethnography**

I have chosen ethnography as a methodology because of its link to the study of culture. I wanted to gain an understanding of the culture in a DID using social constructionism and interpretivism (Creswell, 2007). Interpretivism searches for patterns of meaning, describes meanings, tries to understand the participant’s views and this research is carried out in natural contexts (Gephart, 1999). Constructionism is a form of interpretive research (Gephart, 1999), which states that meaning is not discovered but constructed (Crotty, 2003). In our daily lives the interpretation of our actions by others determines the outcome. Much depends on the way we are perceived and represented by others. In this we have little control on the meaning placed on our actions, and to some extent, our reputation and the way we are seen. This is due to the way in which our society and others make judgements about us and is the key to

constructionism, where meanings and judgements are part of the social group we are part of (Gergen, 1999). I did not want to merely describe the culture as in narrative research or case study research but rather seek to interpret my findings and try to understand the basis of the culture (Creswell, 2007). Ethnography employs several research methods, which link findings together (O'Reilly, 2005). These research methods include observation, interviews, focus groups, and studying artefacts and documents.

Ethnography has its roots in both British social anthropology, where researchers went out to study foreign cultures and in American Sociology (from the Chicago school) which used observation to explore groups on the margins of urban industrial society. The task of these two distinct groups was the same, that of cultural description (Brewer, 2000). Since then ethnography has developed and moved into other spheres such as education, health care and social work. In many respects ethnography is really the most basic form of social research; it bears a close resemblance to the ways in which we make sense of the world around us (Hammersley and Atkinson, 1991). "Ethnography is the art and science of describing a group or culture" (Fetterman, 1989: 11). The study of a culture looks at the way in which people interact and behave when they are part of a community. This community can be a work setting such as a DID.

Ethnography involves the study of a particular social group or culture in naturally occurring settings (McGarry, 2007; Hobbs and May, 1993). Spradley (1979) maintains that the aim of ethnographic research is to gain an understanding of the culture from the point of view of the members of this community. Hobbs and May (1993) concur with this saying that ethnography is a way of telling it like it is, describing the culture observed and looking at the social world being studied as seen from the inside. However Davies (1999) argues that the researcher's understanding of the culture forms the basis of the findings, which come from the information provided by informants. Denzin (1997) agrees with this point saying that "there can never be a final representation of what was meant or said – only different textual representations of different experiences" (p5). There are many interpretations and representations of an experience. The researcher has their own interpretation of an event and the participants may have a different interpretation. The researcher attempts to uncover the participants' interpretation and draw their own conclusion about the event using the many versions that exist to try to make sense of the experience.

In order to document their findings the researcher needs to become part of the culture being studied to gain understanding and insight. The researcher needs to have direct and sustained contact with those being researched within their cultural setting. This involves watching what happens, listening to what is said and asking questions (O'Reilly, 2005). Hammersley and Atkinson (1991) advocate the study of a culture in its natural state, as undisturbed by the researcher as possible. Ethnography should also be carried out over a period of time in order to reduce the impact of the researcher's presence on the situation being studied. "People can sustain an act or maintain their best image only so long" (Wolcott, 199: 49). The researcher's presence may alter behaviour for a short period of time, but this will only continue for a while as 'real' behaviour re-emerges. Nieswiadomy (2002) suggests an adjustment period is needed in order for behaviour to return to normal as people can only maintain an act for a short while.

This study of the culture in a DID explored how DRs made decisions and behaved, and looked at whether this culture is the source of human behaviour or the result of it (Crotty, 2005). Within a cultural setting meanings and actions are based on the meanings and actions of others. These can be modified through observations of and further interactions with others (Crotty, 2005). This can be positive; for example DRs may learn how to deal with difficult patients by observing their colleagues, or it could be negative; for example DRs may follow the example of a colleague in being rude or unhelpful to a referring clinician. This is an example of situated learning (Lave and Wenger, 1991), where members of a community of practice learn from one another in practice about their professional role. This study looked at how DRs interacted with one another, with other health care professionals, with students and with patients. In order to be understood people try to make their actions meaningful to others (Ellen, 1984).

The heart of ethnography is the 'lived order', the way in which members of a group construct, enact, do and inhabit their daily world (Allen, 2004). Ethnography utilises three main research methods; observation, interviews and the study of written documents (Brewer, 2000; Hammersley and Atkinson, 1991). Ethnography is iterative-inductive research, and is an ongoing simultaneous process of theory building, testing and re-building (O'Reilly, 2005).

Ethnography is usually fluid and flexible; a reflexive process with a broad topic and some guiding questions (O'Reilly, 2005).

### **Ethics and ethical issues**

Ethics in research involves the application of ethical principles which include the way in which the research is designed and conducted. The main principle is that participants should not be harmed as a result of participating in the research (Bowling, 2002). All participants should give informed consent in order to participate and this consent should be written (Bowling, 2002). Ethical approval must be sought for all studies using human subjects which take part within the NHS.

Ethical approval was needed for this study from the University of Salford Research Governance and Ethics Committee, the local research ethics committee (LREC) and the research and development committee (R&D) at the NHS Trust where the study took place. Ethical approval for the study was finalised in May 2008.

### **Access to the field**

Long et al. (2008) state that it is not easy to gain access to a hospital for research purposes. Because of my position as a radiography lecturer at the university I was fortunate to be on first name terms with all of the radiology managers in the region that provided clinical placements for diagnostic radiography students. The manager of one of the trusts volunteered to host me and was very interested in my study. It was therefore relatively easy for me to gain access to the DID. Allott and Robb (1998) cite this as a distinct advantage of doing research in your own area of practice.

However, because of the way in which I gained access to the field I was aware of coercion and made every effort to ensure that participants made an informed decision about taking part in the research and did not feel obligated to do so because the manager had given permission for me to work in the DID.

### **Informed consent**

Before the study I had to resolve the issues of informing the staff about the study and gaining consent. It was important that staff members were not coerced into taking part. Therefore I spoke to all of the staff in the DID at their staff meeting and provided each one of them with a participant information sheet and my contact details. After staff members had time to read about the study they were asked to complete the consent form. Staff members were able to opt out of the study at any time.

The LREC asked me to ensure that patients gave their consent for me to observe them. This was achieved by placing a notice in the patient waiting room in the DID and asking each patient being observed for their permission, this was practiced by other similar studies such as May-Cahal et al. (2004). This did not achieve informed consent for the patients but no patient details formed part of the study as my primary focus was on observing the DRs and their practice. The LREC were satisfied with this level of consent for patients as I was abiding by professional code of conduct with regards to patient information.

### **Ensuring no harm**

Before the commencement of the study I had to decide how I would deal with the observation of mal-practice. It was decided in discussion with the manager of the DID that I would intervene if necessary and that I would report any instances. This was difficult for me as I did not feel that this was my role as a researcher, but as a practicing DR I had to resolve my own conscience and I realised that I could not in all good conscience watch whilst a patient was put at risk. Dixon-Woods (2003) says that "ethical issues about when and how to intervene are not uncommon" (p326), and other writers speak about the dilemma of observing bad practice and if intervention is necessary (Hobbs and May, 1993; McGarry, 2007).

During the study there was a mechanism for staff to withdraw from the study. If a staff member decided to withdraw from the study all data relating to them would also be removed from the study.

### **Confidentiality**

The names of staff were not used during the study and staff members were referred to by their profession or title, e.g. nurse, DR, administrative assistant. Each member of staff was also numbered, e.g. DR 1, nurse 2, student DR 4.

It was decided, in discussion with the manager of the DID that I should wear a DR's uniform for the duration of the observation. It was felt that this would be less intimidating for both staff and patients and I would fade into the background more easily. Coffey (1999) says that the researcher should have an acceptable appearance which includes dress, demeanour and speech. Hammersley and Atkinson (1991) agree with this and think that the personal appearance and impression created by the researcher can influence data collection.

I was prepared for staff members to be un-co-operative and in fact at the initial meeting with the staff some of the DRs felt that I might be there to spy on them, a finding which Roper and Shapira (2000) shared. I dealt with this by explaining the reason for my presence and by showing them what I would be doing and trying to ensure that DRs were comfortable with my presence each time I carried out a period of observation. Dixon-Woods (2000) also warns against hostile staff members when observation is part of the research. In reality I did not experience any hostility from DRs whilst I was in the DID, in fact I was welcomed into the team fairly quickly and none of DRs appeared to be worried about my presence.

I ensured that all staff had the opportunity for support should any element of the study cause them distress. This support was provided by the Occupational health department at the NHS Trust where the study was carried out.

### **Situational ethics**

I decided to record my observational data in a notebook which I took with me into the DID. I did leave the notebook and the work surface in the DID when I went into the X-ray rooms, when I did this staff members would have been able to read my notes. In doing this I hoped that staff would realise that I had nothing to hide from them. Assigning numbers to staff members protected their identity. The numbering system was used for the whole study.

### **Methods**

In order to study the culture three main research methods were used; 1) observation within the DID to identify issues 2) interviews with staff members from the DID to further explore the issues highlighted by the observations, and 3) examination of documents used in the DID. The research was carried out in one DID in a medium sized acute NHS teaching hospital by one researcher over a period of seven months. The purpose of the research was not to seek generalisable results but to gain understanding and meanings about the culture in which DRs work (Creswell, 2007).

Kennedy (1999) advocates this type of research using observation in a practice based profession as it allows for the collection of rich data, "observation helps to make sense of the world around us and guides our decisions and actions" (p56). In a profession such as diagnostic radiography there are many complex actions and interactions which can be explored through observation. Ethnography can illuminate hitherto covert patterns of behaviour and decision making (Kennedy, 1999). It is very difficult to explain how professionals behave or why they make certain decisions without seeing these in context. Ethnographic research helps to contextualise behaviour and decision making; it seeks to understand people's actions and their experiences of the world through observing the participants in their natural settings (McGarry, 2007).

## Observation

The study commenced with a prolonged period of observation within a DID. At the beginning of the observation I started with an initial mapping of the DID (Hodgson, 2002). O'Reilly (2005) suggests that a plan or description of the field (in this case the DID) assists in description of the culture. May-Chahal et al. (2004) and Wolf (1988) provide floor plans of the departments/wards in which they carried out their research which can be referred to by the reader to gain an understanding of the location of different events described in the research. The space and place is an important part of the data as it helps to contextualise the findings.

Observation involves all of the senses. Observation involves sound, movement, touch, and smell. Edvardsson and Street (2007) argue for a "sensate field researcher" (p25) who is able to "accurately document and reflect on the use of sensate material" (p30). Using this form of sensate observation allowed me to reflect on how the body is central to any care environment.

The observations for this study were carried out for one day per week. This day was altered each week and I also spent some time observing during the out of hours period (after 5pm), this was in order to observe the DID in its natural state. I took on the role of 'observer as participant' from the four researcher roles in observation outlined by Gold (1958). I considered being a participant observer, but discounted it for this study as I felt that if I was working as a radiographer I may miss out on interactions between staff as I could be alone in an X-ray room imaging patients.

During the period of observation I took field notes in a small notebook. I used these field notes to record my observations and also my own thoughts and feelings about what was going on. Allan (2006) says that the researcher's thoughts and feelings are also important data. In my notes I differentiated between my actual observations and my thoughts on those observations. I felt that it was important to record how I felt about what I had observed. When I typed the observations I used italics to represent my feelings.

The field notes that I took were personal to me and I chose what to record (Coffey, 1999). This involved my decisions about what was significant (Agar, 1980; Anspach and Mizrachi, 2006). Abbott and Sapsford (1997) state that the interpretations, values and interests of the researcher are a central part of the research. My ideas obviously directed where I observed and what I observed.

Observation prompts the researcher to consider what it means to be a part of the group being studied (Allen, 2004). During observation I had to balance the dual roles of professional and researcher. During the whole period of observation I was aware that my insider status could contribute to me missing out on important information (Styles, 1979), this was because I would not necessarily see something as strange or unfamiliar and record this in my notes. I needed to fight familiarity when carrying out my observations and look at the environment with a sense of strangeness (Coffey, 1999). I needed to try to see the DID as through the eye of an outsider, which is often termed the etic perspective (Fetterman, 1989). I had to try to view the environment from a different perspective (Cudmore and Sondermeyer, 2007). I needed to be aware of over familiarisation (Bonner and Tolhurst, 2002), so every day I endeavoured to look around the department for something new that I hadn't seen before or written about. This way I tried to keep my observations fresh and tried to see the environment in a new light.

It was difficult at first to adjust to being in the department but not working there. I had a feeling of guilt about not having a clinical role and being able to assist the radiographers. This was particularly true when the department was busy and all I wanted to do was to take the next X-ray request form and X-ray the patient. Rudge (1995) also highlights this tension and talks about the ethics of assisting in the practice area when your role there is to be a researcher and to observe.

The observation took place over a period of four months. I hoped that over this period of time I began to fade into the background and participants were able to behave as they would if I were not present (Ellen, 1984). Some of the DRs forgot that I was there, whilst others did not appear to consider the reason for my presence and treated me as a member of the team. In studies of this nature the "Halo effect" often occurs (Asch, 1946) where participants being observed want to be seen in a favourable light. Other writers describe the "Hawthorne effect" (Bowling, 2002; Vehmas, 1997) where participants are aware of being observed and alter their behaviour. Some of the radiographers engaged me in the team, and spoke to me frequently, whereas others were quite happy to ignore me. However, after a week of my period of



observation many of the staff members included me in the team and admitted to forgetting why I was actually there. This reinforced my understanding that over a period of time the researcher will begin to fade into the background and participants will behave as they would if the researcher were not present. Ellen (1984) says that this is true after a short period of time and Bowling (2002) says that the "Hawthorne effect" fades over time. It is important to acknowledge that it is not possible to be completely overt; people may forget that the researcher is present and it is not always easy to explain fully the nature of the research (O'Reilly, 2005). It is difficult to balance the need to be open and honest with the need to fit in and become unobtrusive.

My decision to wear uniform helped me to integrate into the department. However, this prompted thoughts about how I felt to be wearing uniform and yet not being involved in the care and imaging of patients. As professionals the wearing of uniform is a powerful statement and it helps us to take on our professional role and persona. I struggled with the fact that I was dressed as a radiographer but was not 'being' a radiographer. This is a concept referred to by Cudmore and Sondermeyer (2007) as being there but not being there. It has been argued that without true immersion in the culture the researcher cannot provide an authentic account (Allen, 2004). Therefore I spent the whole of each day of the observation with staff including eating lunch and taking tea breaks in the staff room. I felt that this helped me to become integrated into the team and recognised as a part of the staff group.

I was able to use both structured and unstructured observations during this time. Structured observations were carried out through the observation of one particular DR over a period of one complete shift using an observation chart which documented their movements, actions and interactions (May-Chahal et al., 2004). "Observation is pivotal to the way in which skills are passed on and things are known" (Grimshaw and Ravetz, 2005: 74). Unstructured observations were carried out in significant areas of the DID, which included the main viewing area, the staff room and the patient waiting areas. When observing in these locations I made field notes about actions, behaviours and interactions which were observed. I chose the locations for the unstructured observations after my initial survey of the DID where I tried to determine the main areas of the department where interactions between staff took place and areas which provided me with useful and meaningful data. I was able to observe a cross section of the staff in the DID.

Another challenge was being able to fit in, in order to cause as little disruption as possible (Bonner and Tolhurst, 2002). I intended to become a familiar part of the work setting within the DID in order that staff members continued to work as normal. Coffey (1999) encourages carving out a space to be, a location that allows for observation but does not intrude on events. To this end I selected places to stand that were as unobtrusive as possible. This often involved standing in a corner in the viewing area or behind the lead glass screen of an X-ray room where I could see what was going on but I wasn't in the way of the DRs and did not interrupt their work flow.

During the period of observation I learnt my own style. When I started I found it difficult to decide what to record and how to record it. I developed a note form with my own abbreviations which I typed up as soon as I reached home when the work was still fresh in my mind. I used the drive home to reflect on my day and did a lot of thinking in the car. I was keen to formally record the data as soon as possible after the event to reduce the chances of inaccuracy.

## **Interviews**

Interviews with key informants were used following the observations to explore issues further. I was able to interview a cross-section of staff from the DID. May-Chahal et al. (2004) were able to gather useful information from a variety of staff groups in the department where they carried out their study, and this included domestic staff. For my study 10 interviews were carried out with key informants. The key informants were identified during the observations and I selected these people, this was a purposive sample in order to gather meaningful data (Bowling, 2002).

The interviews were semi-structured and explored further the issues highlighted by the observations (Coffey, 1999). I was able to seek clarification about issues from the participant's perspective. Different staff groups were chosen with the intention of choosing some leaders and some followers. Leadership has an influence upon the culture (Wolf, 1988),

and I wanted to see if leaders and followers have different perspectives. There is a structured hierarchy within the DID with different staff being responsible for and leading teams of people.

The questions used during the interviews were open and exploratory questions. These questions were based on the themes extracted from the observations and also explored further some of issues uncovered by the literature review. The interviews were recorded onto a digital recording device and transcribed verbatim. The data produced was contextualised and I began to look at issues and events from the insider's or emic perspective (Fetterman, 1989). Validation of findings can be done by examining all of the data from a study to test the findings. Results can be confirmed by using data from different sources and this helps to give authority to the findings (Brewer, 2000). However, it is important to acknowledge that the final ethnographic report is not 'the truth', it is the researcher's representation, the researcher's 'voice' (Allan, 2006). I was able to look for patterns of behaviour, action and interpretation (Fetterman, 1989; Hodgson, 2002).

### **Examination of documents**

Documents kept and used within the DID were also examined. "Documents are produced by humankind in socially organised circumstances" (Prior 2003: 4). The documents examined included policies and procedures as well as minutes of meetings, for example; the departmental examination protocols, infection control procedures, the DID moving and handling policy, the major incident procedure, minutes of superintendents meetings, minutes of staff meetings and memos placed on notice boards. The examination of documents provided further insights into and information about the organisational structure and culture of the DID.

Documents can be full of concepts, assumptions and ideas (Prior, 2003). The examination of documents highlighted issues about how the structure and culture of the DID affects staff members (Allen, 2000). The documents also uncover hierarchy and power relationships within the DID (Becker et al., 1961; Lofland and Lofland, 1984). Davies (1999) says that documents should be examined in terms of text/discourse, interaction and context. I was able to look at how documents are written and how information is recorded in the DID, both formal and informal. The language used within a culture is often shared professional terminology which is only understood by those who are part of the culture. Beals et al. (1977) talk about the artefacts used within a culture and documents are some of these artefacts. In studying these I was able to gain further information about the DID and the culture of information giving within the DID.

### **Initial findings**

1. Time pressures versus patient care: DRs often work in a pressured environment where there is a high demand for the service and where there are many targets to do with throughput of patients and patient waiting times. DRs are constantly weighing this up against the care they can give to their patients within the short time they have to spend with them.
2. Systems of work – 'the way we do things around here': DRs work in a systematic and structured way and have a set way of doing things. They work to policies, protocols and procedures.
3. Use of 'black humour': DRs use humour as a coping mechanism for dealing with stressful and difficult situations.
4. Team working: DRs work together well in a team and use verbal and non-verbal communication to communicate their feelings.
5. Relationships with patients: DRs spend a very short period of time with patients. DRs make a quick initial assessment of the patient's capabilities and needs. The short interaction that occurs is very goal focussed, i.e. it is concerned with the production of a good quality diagnostic image.
6. Extended role: There are many opportunities for DRs to take on extended roles and their relationship with radiologists has an effect on the opportunities available.
7. Use of ionising radiation: DRs cite this as a reason for much misunderstanding between DRs and professional groups. DRs can use their knowledge to exercise power.

8. *Professional socialisation:* The way in which DRs learn to become a DR was evident, particularly when observing students. DRs learn how to behave by observing and copying the behaviour of other DRs.
9. *Behaviour in different parts of the department:* There were different patterns of behaviour in front of patients and behind the scenes.
10. *Keeping patients waiting:* DRs did not like to do this and were highly anxious about it in all areas of the DID. However, they tried not to mention the waiting time to patients as they tried to avoid confrontation.
11. *Working with equipment/computers:* The computers and X-ray equipment are as much part of the culture as the people.
12. *Discussing work with others:* DRs discuss their work; patients, images, image requests and patient positioning with their colleagues.
13. *Story telling:* This is very much a part of staff room and staff area culture.
14. *Referring to patients by the name of their examination:* DRs tend to say things like, 'there's a foot waiting outside', or 'there's a spine on a trolley waiting'. This is seen to be perfectly normal.
15. *Visible product:* DRs tend to be concerned about the image they have produced. Perhaps this is because the result of their interaction with the patient is there for all to see.

These are only initial findings from the data and a more thorough analysis will be carried out.

### **Data analysis**

Data analysis will be carried out using a thematic analysis. The observation, interview and document analysis data will be compared to look for common themes.

### **Limitations**

There are some limitations with my choice of methodology and methods. The main limitation is that this was a small study, carried out in just one DID. Sampling within the DID was important, i.e. when and where to observe, who and when to interview and which documents to review (Hammersley and Atkinson, 1991). The choices I made and sampling will ultimately influence the data gathered. It was not possible for me to observe all of the workings of the DID and I may have missed out on important interactions. I will only be able to work with the data that I am able to gather. Participants could have been un-co-operative (Lofland and Lofland, 1984), or they may have chosen to withhold information (Roberts, 2007). Participants may have maintained an 'act' whilst I was observing, not providing a true picture of the situation (Wolcott, 1999). Participants may have felt threatened by the situation and not felt able to reveal information (Allen, 2004).

### **References**

- Abbott, P and Sapsford, R (Eds.) (1997) *Research into Practice – A Reader*, OU Press, Buckingham
- Adams, J and Smith, T (2003) 'Qualitative methods in radiography research: a proposed framework', *Radiography*, 9 (3), Aug. 2003. pp. 193-199.
- Agar, M H (1980) *The Professional Stranger – An Informal Introduction to Ethnography*. Academic Press, London.
- Allan, H T (2006) 'Using participant observation to immerse oneself in the field: The relevance and importance of ethnography for illuminating the role of emotions in nursing practice', *Journal of Research in Nursing*, 2006: 11, pp. 397-407
- Allen, D (2000) 'Doing Occupational Demarcation: The Boundary Work of Nurse Managers in a District General Hospital', *Journal of Contemporary Ethnography*, 29 (3), June 2000, pp.326-356
- Allen, D (2004) 'Ethnomethodological insights into insider-outsider relationships in nursing ethnographies of healthcare settings', *Nursing Inquiry* 2004, 11(1), pp.14-24.

- Allott, M and Robb, M (Eds)(1998) *Understanding Health and Social Care – An Introductory Reader*, Sage, London.
- Anspach, R R, Mizrahi, N (2006) 'The field worker's fields: ethics, ethnography and medical sociology', *Sociology of Health and Illness*, 28(6) 2006, pp 713-731.
- Asch, S E (1946) 'Forming impressions of personality', *Journal of Abnormal and Social Psychology*, 41, pp.258-290.
- Beals, R L, Hoijer, H and Beals, A R (1977) *An Introduction to Anthropology* (5<sup>th</sup> Ed.) Collier Macmillan Publishers, London.
- Becker, H, Geer, B, Hughes, E C & Strauss, A L (1961) *Boys in White – Student Culture in Medical School*, Transaction Publishers, New Brunswick.
- Bonner, A and Tolhurst, G (2002) 'Insider-outsider perspectives of participant observation', *Nurse Researcher* 9 (4) pp.7-19.
- Bowling, A (2002) *Research methods in health – investigating health and health services* (2<sup>nd</sup> Ed.) Open University Press, Maidenhead.
- Brewer, J D (2000) *Ethnography*, Open University Press, Buckingham.
- Coffey, A (1999) *The Ethnographic Self*, Sage, London
- Creswell, J W (2007) *Qualitative Inquiry and Research design – Choosing among five approaches* (2<sup>nd</sup> ed.) Sage, London.
- Crotty, M (2005) *The Foundations of Social Research – Meaning and Perspective in the Research Process*, Sage, London.
- Cudmore, H, and Sondermeyer J (2007) 'Through the looking glass: being a critical ethnographer in a familiar nursing context', *Nurse Researcher*, 2007, 14:3, pp25-35.
- Davies, C A (1999) *Reflexive Ethnography – A Guide to researching selves and others*, Routledge, London.
- Denzin, N (1997) *Interpretive Ethnography – Ethnographic Practices for the 21<sup>st</sup> Century*, Sage, Thousand Oaks.
- Dixon-Woods, M (2003) 'What can ethnography do for quality and safety in health care?', *Quality and Safety in Health Care*, Oct.2003, 12, 15, pp.326-327.
- Edvardsson, D and Street A (2007) 'Sense or no-sense: The nurse as embodied ethnographer', *International Journal of Nursing Practice* 2007, 13, pp. 24-32.
- Ellen, R F (Ed.) (1984) *Ethnographic Research – A Guide to General Conduct*, Academic Press, London.
- Fetterman, D J (1989) *Ethnography – Step by Step*, Sage, California.
- Gephart, R (1999) 'Paradigms and Research Methods', *Research Methods Forum*, Vol. 4, Summer 1999, University of Alberta.
- Gergen, K J (1999) *An invitation to Social Construction*, Sage, London.
- Gold, R L (1958) 'Roles in Sociological fieldwork', *Social forces*, Vol. 36, pp217-223.
- Grimshaw, A and Ravetz (Eds.) (2005) *Visualising Anthropology*, Intellect Books, Bristol.
- Hammersley, M and Atkinson, P (1991) *Ethnographic Principles in Practice*, Routledge, London.
- Hobbs, D & May, T (Eds.) (1993) *Interpreting the field – Accounts of ethnography*, Oxford University Press, Oxford.
- Hodgson, I (2002) 'Engaging with cultures – Reflections on entering the ethnographic field', *Nurse Researcher*, 9 (1), pp. 41-51.
- Kennedy, C (1999) 'Participant Observation as a research tool in a practice based profession', *Nurse Researcher*, 7 (1), pp. 56-65.
- Lave, J and Wenger, E (1991) *Situated learning. Legitimate peripheral participation*, Cambridge University Press, Cambridge.
- Lofland, J and Lofland, L H (1984) *Analyzing Social Settings – A Guide to Qualitative Observation and Analysis*, (2<sup>nd</sup> ed.) Wadsworth, California.

- Long, D, Hunter, C L and Van Der Geest, S (2008) 'When the field is a ward or clinic: Hospital ethnography', *Anthropology and Medicine*, 15(2) August 2008, pp71-78.
- May-Chahal, C, Hicks, S and Tomlinson, J (2004) *The relationship between child death and child maltreatment – A Research study on the attribution of cause of death in hospital settings*. NSPCC, London.
- McGarry, J (2007) 'Nursing relationships in ethnographic research: what of rapport', *Nurse Researcher*, 2007, 14 (3) pp.7-14.
- Ng, K C, and White, P (2005) 'Qualitative research design and approaches in radiography', *Radiography*, 11 (3), August 2005, pp.217-225.
- Nieswiadomy, R (2002) *Foundations of Nursing Research*, (4<sup>th</sup> Ed.) Prentice Hall, Upper Saddle River, USA.
- O'Reilly, K (2005) *Ethnographic Methods*. Routledge, London.
- Prior, L (2003) *Using Documents in Social Research*, Sage, London.
- Roberts, D (2007) 'Ethnography and staying in your own nest', *Nurse Researcher* 2007, 14(3) pp15-24.
- Roper, J M and Shapira J (2000) *Ethnography in Nursing Research*, Sage, London.
- Rudge, T (1995) 'Response: Insider ethnography: researching nursing from within', *Nursing Inquiry*, 1995; 2; 58.
- Spradley, J P (1979) *The ethnographic interview*, Holt, Rinehart and Winston, London.
- Styles, J (1979) 'Outside/insider: Researching Gay Baths', *Urban Life* 8(2), pp135-152.
- Vehmas, T (1997) 'Hawthorne effect: shortening of fluoroscopy times during radiation measurement studies', *British Journal of Radiology*, 70: 838, pp.1053-1055.
- Wolcott, H F (1999) *Ethnography – A Way of Seeing*, Altamira Press, Oxford.
- Wolf, Z R (1988) *Nurses' work, the sacred and the profane*, University of Pennsylvania Press, Philadelphia.

# The significance and co-dependency of 'Robustness' from the perspective of the Granting Authority in a PFI project environment

**Gerald Sundaraj, Prof. David Eaton**, School of Built Environment, University of Salford

## **Abstract**

The usage of Public Private Partnership (PPP) in the construction environment at the project conceptual stage has been well established globally. In the UK, the practice of using the PPP philosophy began as early as 1992. As there are many different methodologies of PPP, the Private Finance Initiative (PFI) approach has been widely accepted in the UK and other parts of Europe. The PFI model is particularly prominent compared to other PPP models due to the element of Value for Money (VfM). VfM can be achieved provided a balanced Risk Transfer assessment is carried out by the parties involved. PFI needs to provide elements of Affordability in order to sustain it throughout the duration of the contract, usually 25-30 years. With the investment of time and cost in a PFI project, the element of Robustness in a PFI model will lead to it being reproduced to fit similar or related projects, thus providing sustainability. The research is focused on exploring the co-dependency between Risk Transfer-VfM-Affordability-Robustness in a PPP/PFI model. Based on the co-dependency, the study will probe to explore a balance of the four elements in the PPP/PFI model. This research will be approached from the perspective of the Granting Authority.

## **1.0 Introduction**

This paper is presented as part of an ongoing PhD research undertaken at the University of Salford. The study focuses on the Public Private Partnership (PPP) in general and the Private Finance Initiative (PFI) specifically. A blurry demarcation exists between the two terms as PFI is more widely used in the UK. The research will explore the significance and co-dependency of 'Robustness' within a PFI project environment and from the perspective of the Granting Authority. In order to provide some direction to the research, an aim, objectives and research questions have been developed, however they are subject to refinement as the research matures over time.

### *Research Aim:*

To determine the significance and co-dependency of 'Robustness' in a PPP (PFI) project environment from the perspective of the Granting Authority.

### *Research Objectives:*

- i. To determine the significance of Robustness in a PPP(PFI) project environment;*
- ii. To determine the interdependency and co-dependency of the Critical Success Factors (CSFs);*
- iii. To determine a sustainable CSFs model/framework.*

### *Research Questions:*

- What is the link between Risk Transfer Assessment, Value for Money, Affordability and Robustness?*
- How do these four elements affect each other individually & collectively to provide a balance in a PPP/PFI model?*
- How can these four CSFs provide sustainability to the PFI model in a PFI project environment?*

The methodology component of this research is currently being explored and is not covered as part of this paper. In order to obtain a more clear understanding of PPP and PFI, it is important to define the two terms which is the core component of this study.

## **2.0 Defining PFI & PPP**

Over the last the two centuries, the construction industry has experienced several transformations. Events such as World Wars, recessions, epidemics, oil crises and the most recent financial crisis have moulded the global construction industry. Nevertheless these changes were vital to meet the demands of the challenges faced. However the rate at which the challenges had occurred were far greater than that of the changes, but these variations were important to drive the construction industry forward. In the UK itself, the construction industry had experienced key transformation moments which have led to the industry as it is today. Changes brought during the Thatcher administration (1979-1990), privatization had laid the foundation for the private sector to play a prominent role in the delivery of public services for years to come. The launch of the Private Finance Initiative (PFI) in 1992 and the Public Private Partnership (PPP) program in 1997 (Spackman, 2002) had changed tremendously the way the public and private sector would be operating within a project environment as well as the construction industry as a whole. These changes widened the construction fraternity, introduced new contract management and accounting practices, increased business opportunities and interaction between the public and private sectors.

The Private Finance Initiative (PFI) was first introduced in the United Kingdom in 1992 (Payne, 1997) to further engage the participation of the private sector in the delivery of public services. This innovative procurement concept, born in the UK, attracted Governments from various countries as public financial constraint was taking a toll globally and was coupled with raising development expenditures. Besides the option of the 'off-balance sheet' accounting, PFI provided numerous benefits to the public sector and private sector. To appreciate the PFI concept even more, several definitions are explored to obtain a wide perspective of the industry.

Dick *et al*, (1996:4)

"PFI is typically based around a concession agreement."

Akintoye *et al*, (1999:81)

"The Private Finance Initiative (PFI) involves a process where private sector organisations, design, build, finance and operate assets which deliver a service to public sector clients"

HM Treasury, (2000:47)

"The public sector contracts to purchase quality services, with defined outputs, on a long-term basis from the private sector, and including maintaining or constructing the necessary infrastructure. The term also covers financially free-standing projects where the private sector supplier designs, builds, finances and then operates an asset and covers the costs entirely through direct charges on the private users of the asset, with the public sector involvement limited to enabling the project to go ahead through assistance with planning, licensing and other statutory procedures."

De Lemos *et al*, (2003:59)

"PFI is defined as an integrative, holistic, social and economic developing policy of the UK Government. It integrates a hierarchy of four levels of concepts: government ideology, principles, practices and tools. In essence it is the provision of a public service funded by private capital."

Ball and King, (2006:36)

"The Private Finance Initiative (PFI) is, however, a type of PPP with a highly prescriptive legal framework and relates to the provision of capital assets for the public service."

Dixon and Pottinger, (2006:479)

"...the PFI is a means of procuring public sector infrastructure and services using private sector finance and expertise."

Eaton, (2005:1)

"...a legal framework for concessions in the UK to encourage private capital investment in the construction industry."

Eaton *et al*, (2007:1472)

"The PFI is therefore a generic classifier for all types of 'construction PPP'."

In summary, PFI can be defined as a partnership between the public and private sectors, implemented through a series of procurement tools and regulatory frameworks to improve socio-economic development by capitalising on the private sector's expertise and finance for the delivery of public services.

The initial progression of PFI had been relatively slow both within central government and local authorities (Harding *et al*, 2000). This was caused by several factors which have been extensively identified (Ezulike *et al.*, 1997; Owen and Merna, 1997). These barriers have been in existence due to the 'change' factor which was required to push the PFI agenda. The feeling of operating in unfamiliar grounds by both the public and private sector raised numerous questions which needed immediate action, notably from the public sector. The response was delivered via a number of political initiatives: the Public Private Partnership Program (4Ps), 1996; the Bates Review I, 1997; the Treasury Taskforce, 1997; the Bates Review II, 1999; the Gershon Review, 1999; the establishment of the Office of Government Commerce (OGC), 1999 and Partnership UK (PUK), 2000 (Eaton, 2005). There was a constant drive from both the public and private sector to see the success of the PFI policy. Nevertheless a shortfall in implementing PFI projects particularly in the health sector was present due to the complexity, flexibility, quality and cost (McKee *et al.*, 2006).

The interaction between the private sector and the public sector has been well established in the construction industry globally (Ghobadian *et al*, 2004). Examples of collaboration between the two sectors dates back to the 1950s in Hong Kong with the construction of a privatised vehicle tunnel (Grimsey & Lewis, 2002). Grimsey and Lewis (*ibid*) goes further to indicate that Manod, (1982) stated that the first concession was granted in 1782 in France. The US federal government has been using PPP as a tool to stimulate private sector investment since the 1960s (Pongsiri, 2002). This reflects the acceptance of PPP, not just in the UK but globally as an alternative procurement mechanism to the traditional procurement system. To understand the Public Private Partnership concept, several definitions are explored from various organisations and journal papers before summarising a personal definition.

HM Treasury, (2000:10)

"Public Private Partnerships bring public and private sectors together in a long term partnership for mutual benefit"

Grimsey & Lewis, (2002:108)

"PPP is defined as agreements where public sector bodies enter into long term contractual agreements with private sector entities for the construction or management of public sector infrastructure facilities by the private sector entity, or the provision of services (using infrastructure facilities) by the private sector entity to the community on behalf of a public sector entity"

Klijn & Teisman, (2003:137)

"PPPs can be defined as co-operation between public and private actors with a durable character in which actors develop mutual products and/or services and in which risk, costs and benefits are shared. These are based on the idea of mutual added value"

Akintoye, (2009:124)

"PPP can be described as a contractual agreement of shared ownership between a public agency and a private company, whereby, as partners, they pool resources together and share risks and rewards, to create efficiency in production and provision of public or private goods"

Based on the various definitions, Public Private Partnership can be defined as a procurement mechanism which is contractually binding between the public and private entity with specific mutual gains and risks from the partnership to provide construction and management services over a period of time, through which an asset, a service or both is delivered to provide public services. The definition indicates that the partnership is contractually obligatory which requires the support of an effective regulatory framework (Pongsiri, 2002). The regulatory framework



will provide an avenue to define the roles and responsibility of the parties as well as preserving their interests and investment in the partnership. As it is a partnership approach, an understanding of each others objectives is crucial to the success of the partnership. The public sector is regarded as a bureaucratic, over regulated and monopolised environment in which public enterprises operate to ensure sustainability (Ghobadian *et al.*, 2004b; HM Treasury, 2000). The private sector on the other hand is driven by competition, deregulation and profit orientated culture (*ibid*). The integration between these sectors towards a mutual benefit will strengthen the partnership which leads to a positive outcome in delivering public services.

### **3.0 The UK Experience**

The change in the Government's approach to the delivery of public services started during the Thatcher administration in 1979 by reversing the nationalisation of key industries such as shipbuilding, vehicle manufacturing, and public services such as electricity, telecommunication and water through the concept of privatisation (Ghobadian *et al.*, 2004). The privatisation practice under the Conservative Party allowed the private sector to participate actively in the delivery of public services, notwithstanding the fact that the private sector has always provided services for the public sector through the conventional procurement method and contracting out; nevertheless privatisation was the catalyst in allowing the private sector to play a leading role in the delivery of public services for much time to come. Privatisation in its optimum state was found to highlight world class construction companies, providing cost and price effectiveness and services orientated to meet customers' needs (HM Treasury, 2000). Based on deficiencies in the implementation of privatisation highlighted by the HM Treasury (*ibid*), it can be concluded that over time, the privatisation exercise provided more information on how privatisation would benefit the Government as well as the end user and that competition was key to innovation which would lead to efficiency and effectiveness and that a regulatory framework was needed to support initiatives by the Government as well as ensuring the consumers a fair deal. With the growing interest in using private financing as a means to implement public services from various government entities, the 'Ryrie Rules' were formulated to assist in filtering proposals that merely tried to prevaricate against expenditure procedures. The 'Ryrie Rules' were retired in the late 1980s as it was seen as a barrier for exploring private sector participation and private financing holistically, as the former was seen to bring efficiency into public service delivery. This concurred with Spackman, (2002) that the public sector, through the off-budget finance element, would be able to weed out inefficiencies.

Construction demand in the early 1990s was rising at the same time as the cost of capital expenditure in the UK. The Government was equally adamant to restrict the Public Sector Borrowing Requirement (PSBR), forcing them to seek alternatives to meet the needs of the increasing construction demand and cost, opening the window of opportunity for private sector financing (Dick, 1996) through the Private Finance Initiative (PFI). The PFI was launched by the UK Government following the announcement by the Chancellor, Norman Lamont (Akintoye, 1999; Ezulike *et al.*, 1997). It was the Government's intention to foster greater participation from the private sector in the delivery of services which were traditionally within the purview of the public sector. Besides the financial constraints, other factors had contributed to the PFI initiative; among them were the modernisation of Government, the birth of EU legislation, the maturing Victorian infrastructure and the growing complexity of financing projects. The exposure provided by the privatisation experience is also seen as a contributing factor for the partnerships between public and private sector to take place (Broadbent & Laughlin, 2003; Pongsiri, 2002). Broadbent and Laughlin (*ibid*) further argued that the shift was based on the fact that the Government had exhausted its options in the privatisation exercise and needed to unequivocally continue the participation of the private sector, thus the PFI initiative was born.

With the launch of the PFI initiative, the UK Government had also introduced the 'universal testing rule' which subjected all projects to be considered under the PFI option before seeking other procurement alternatives. This provided a boost in the number of projects implemented under the flagship of PFI (Spackman, 2002), the reason being the elements of significant risk transfer from the public sector to the private sector and the offer of value for money as the core principles (Fox and Tott, 1999). However, this approach was seen as rather forceful to opt for PFI in order to implement a project. This created an atmosphere of, for a go ahead in projects using PFI, which created a sudden surge of projects for approval that led to backlogs. The three basic forms of PFI based projects are services sold to the public sector as purchaser and user, financially free-standing and joint ventures (Fox and Tott, 1999). With the change in

the political scenario taking place towards the late 1990s, the PFI approach continued to be supported by the opposition even before coming into office but the initiative was to be part of a more holistic cooperation between the public and private sector through the Public Private Partnership (Broadbent and Laughlin, 2003).

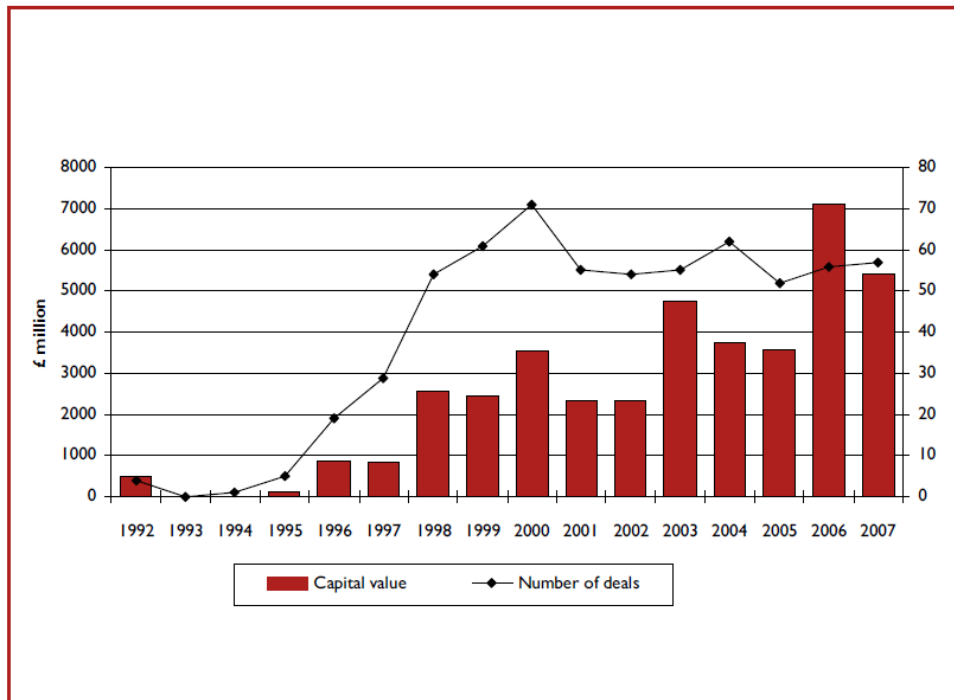
The Public Private Partnership introduced in 1997 by the Labour Government in the UK has been a concept to coordinate and synergise the UK construction industry in becoming more effective and efficient. The usage of Public Private Partnership (PPP) in the construction environment from the project concept stage has been well established globally. PPPs in infrastructure development involve the private sector in the designing, financing, construction, project management, ownership and operation of a public sector utility or service. This form of partnership between the public and private sector is an accepted alternative to the traditional procurement method of public facilities and services. There are several forms of private sector involvement in the public domain, namely service contracts, leasing, joint ventures, concessions and privatisation (Li & Akintoye, 2003). In the UK, the Government has identified eight PPP models for public services and facilities procurement (HM Treasury, 2000:46):

- i. Asset Sales: the sales of surplus public sector assets;
- ii. Wider Market: introduces the skills and finance of the private sector to help with better use of assets (both physical and intellectual) in the public sector;
- iii. Sales of Business: this relates to the sales of shares in state-owned business by flotation or trade sale, with the sale of a minority or majority stake;
- iv. Partnership Companies: introducing private sector ownership into state-owned business, while still preserving public interest and public policy objectives through legislation, regulation, partnership agreements, or retention by Government of a special share;
- v. Private Finance Initiative: the public sector contracts to purchase quality services, with defined outputs, on a long term basis from the private sector, and including maintaining or constructing the necessary infrastructure. The term also covers financially free-standing projects where the private sector supplier design, builds, finances and operates an asset and covers the costs entirely through direct charges on the private users of the asset, with public sector involvement limited to enabling the project to go ahead through assistance with planning, licensing and other statutory procedures;
- vi. Joint Ventures: partnerships in which public and private sector partners pool their assets, finance and expertise under joint management, so as to deliver long term growth in value for both partners;
- vii. Partnership Investments: Partnerships in which the public sector contributes to the funding of investment projects by private sector parties, so as to ensure that the public sector shares in the return generated by these investments;
- viii. Policy Partnerships: arrangements in which private sector individuals or parties are involved in the development or implementation of policy.

The PPP policy of the Labour Government has been a part of their modernisation efforts and to draw out the commercial, financial and management skills of the private sector. The Labour administration brought several key improvements to further invigorate PFI. Recommendations put forth through the Bates Review I, Bates Review II, and the Gershon Review saw the setting up of instruments, such as the Treasury Taskforce, Partnership UK and Office of Government Commerce, needed to drive the PFI and PPP policy forward (HM Treasury, 2000). The most widely used form of PPP has been the PFI approach, primarily because of the value for money reasons (HM Treasury, 2008). Figure 1.0 indicates a rising trend in the number of PFI projects carried out between 1992 and 2007. Through this initiative the public sector shifts its role from a purchaser of assets to a user of services. This transition in role has encompassed the element of innovation in PFI based projects through which the public sector

specifies the output requirements, giving the freedom to the private sector to meet those requirements.

The increasing trend in capital value and number of PFI deals, reflect the growing acceptance and influence of a working model of a partnership between the public and private sectors to provide public services in various sectors. Through the PPP concept, PFI in the UK context has completed 510 projects as of 2007 mainly in education, health, transport, housing, defense, waste management and climate change (HM Treasury, 2008).



Source: HM Treasury, 2008

**Figure 1.0.** Development of PFI over time

The broad overview of the PPP globally and in the UK, shows that the public and private sectors have been continuously engaging each other as a part of an economic cycle. Each party has benefited from one another in some way, notably the private sector which is more profit driven when compared to the public sector, which is focused more on social responsibility. The objectives can be seen as opposites; nevertheless PPP has enabled a convergence of the two components towards achieving their individual goals through a common framework. The tools used in this framework can be seen in a variety of models from concessions, contracting out, joint ventures, PFI and out-sourcing. Limitations in the public sector capital financing have led the PFI model to provide a conduit for financing solutions to be considered by various Governments globally and in the UK. The Private Finance Initiative model has also allowed the various contract formats such as DBFO, BLT, BOT, BOOT, PFI and others to be formally recognized as part of a larger framework which is the PPP.

The partnership between the public and private sector in delivering public services is on a long term basis, usually between 15-30 years (HM Treasury, 2006). Through this hybrid cooperation, the public sector becomes a purchaser of services rather than assets; furthermore the asset delivered by the private sector is based on the output specifications provided by the granting authority in order to achieve the required service outputs. In this arrangement, the private sector bears the risk of upfront capital investment, operational and maintenance cost of the asset as unitary payments are only paid subjected to completion and performance of the asset to deliver the specified outputs (HM Treasury, 2003). Besides VfM and risk transfer, affordability is critical in the delivery of public services through the PFI initiative (Grimsey and Lewis, 2005). Affordability provides a sustainable commitment from the granting authority throughout the duration of the initiative; in this case, on a long term basis. The underpinning factors of a successful PFI initiative, governed by affordability, VfM and risk

transfer has been well documented by several authors (Fourie and Burger, 2000, 2001; Gosling, 2004; Hodge, 2004; Grimsey and Lewis, 2005). De Lemos (2003) claimed that PFI has established a competitive advantage over traditional procurement methods used by the UK Government; nevertheless this advantage is still subjective due to the nature of PFI contracts existing over a 15-30 year period, subjecting it to the whole life cycle factor.

#### 4.0 PFI Life Cycle Cost (LCC)

The relationship of the CSFs within a PFI project environment is explored through the life cycle of a PFI initiative. De Lemos (2002) defined that life cycle of a project as the time span over which the project develops from the very beginning until the end. Comparatively in a PFI project environment context, the life cycle would span from the procurement stage through to delivery and occupation by the end users (Dixon *et al.*, 2005). Since the financial element is inseparable when it comes to construction, thus the life cycle costing of a PFI project environment is equally important and due to the long term nature of this initiative, the growing consideration of life cycle costing can be attributed to the increasing acceptance of PFI as a procurement method (Pasquire and Swaffield, 2002). The close relationship between PFI and LCC is apparent as the latter is used as a management tool as stated by Lemos (2002) in achieving profitability. Subsequently Ball and King (2006) argued that there is little empirical data to support the impact of LCC in PFI projects. A study by Swaffield and McDonald (2008) indicates that Quantity Surveyors acknowledge the importance of LCC to be considered within a PFI project environment as it can lead to an increase in maintenance cost which could affect the overall profit margin of the contractor. While exploring the PFI life cycle, several definitions with regards to this process are defined to provide an understanding of the terms which will be used throughout this section.

Norman, (1990:344)

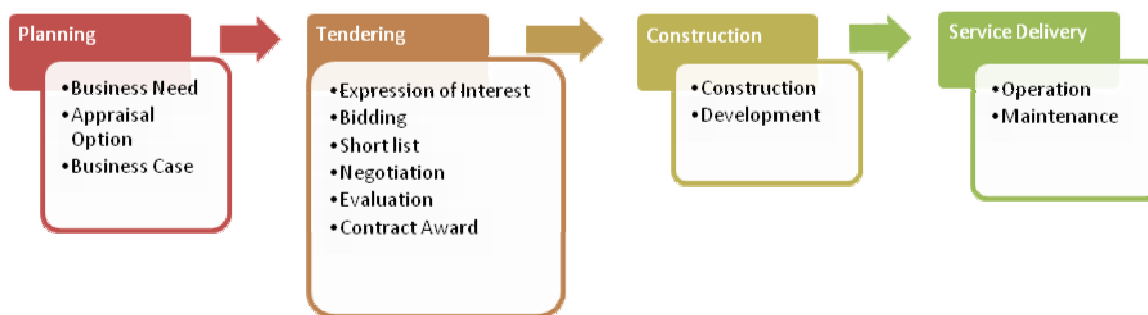
“Life cycle costing is the process of economic analysis that assesses the total cost of investment in and ownership and operation of the system or product to which the life cycle costing analysis is being applied.”

Norman, (1990:344)

“Life cycle cost (LCC) in turn is defined as the total cost of the system or product under study over its complete life or the duration of the period of study, whichever is shorter.”

The association between LCC and the CSFs in a PFI arrangement is self explanatory through the need to achieve an economical and a sustainable financial environment throughout the duration of the contract. VfM, risk transfer and affordability go hand in hand as they are the fundamental objectives in a PFI project and these three can be seen as the processes in the context of life cycle costing.

The life cycle of a PFI can be generally categorised into four phases (Figure 2.0). The progression of a PFI project begins with the planning phase of the project then the tendering phase. This is followed by construction of the asset and finally once completion begins the service delivery phase over specified duration. With respect to PFI schemes usually lasting for a period of 25-30 years, the possibility of retendering the service delivery component to the private sector lies open.



**Figure 2.0** The Phases of a PFI project

The UK being part of the European Union (EU), its procurement of goods and services by the public authorities is governed by the European Union Directives (NAO, 2007). This is a part of an holistic approach to provide a level playing field to contractors in the EU member states. As of 31 January 2006, changes made in the EU Directives have been incorporated into the UK law and among the significant changes is the new procurement procedure of Competitive Dialogue for multifaceted projects, such as PFI schemes (NAO, 2007). The Competitive Dialogue element from the granting authority perspective allows less room for alterations to the scope once the preferred bidder has been appointed. Flexibility within the Competitive Dialogue framework for the contractor in a project environment is provided to obtain clarity for the required output specification by the granting authority. However, the infancy of the Competitive Dialogue raises doubts of its effectiveness (Dorn *et al.*, 2008; Hellowell and Pollock, 2009) and an increase in cost and time that affects VfM and affordability in a project is possible (KPMG, 2008). The comparison between the competitive dialogue and negotiated procedure is provided in Figure 3.0.

The resilience of a PFI contract over such a period of time depends on the VfM, risk transfer and affordability factors. The co-existence of these factors simultaneously in a PFI project environment is crystal clear. Thus there exist interrelationships among these factors which can be categorised as a fourth element. The role of this fourth factor is to provide linkages and equilibrium to the equation to achieve sustainability throughout the whole life cycle of the project. This fourth factor of resilience is robustness. The research will explore the element of robustness to determine the significance and co-dependency within a PFI project environment. It is envisioned that the Critical Success Factors (CSFs) for a PFI project are mainly value for money, risk transfer, affordability and robustness.

## **5.0 Critical Success Factors (CSFs)**

Rockart (1982) defined CSFs as 'the few key areas of activity in which favourable results are absolutely necessary to achieve goals'. This tool is applied through a qualitative process, mainly interviews, to determine the specific areas of activity across the managerial levels that affect performance. The interviews are targeted at key personnel at different hierarchical levels within the organisation. Boynton and Zmud (1984:24) identified several strengths and weaknesses in determining the CSFs as:

### *Strength:*

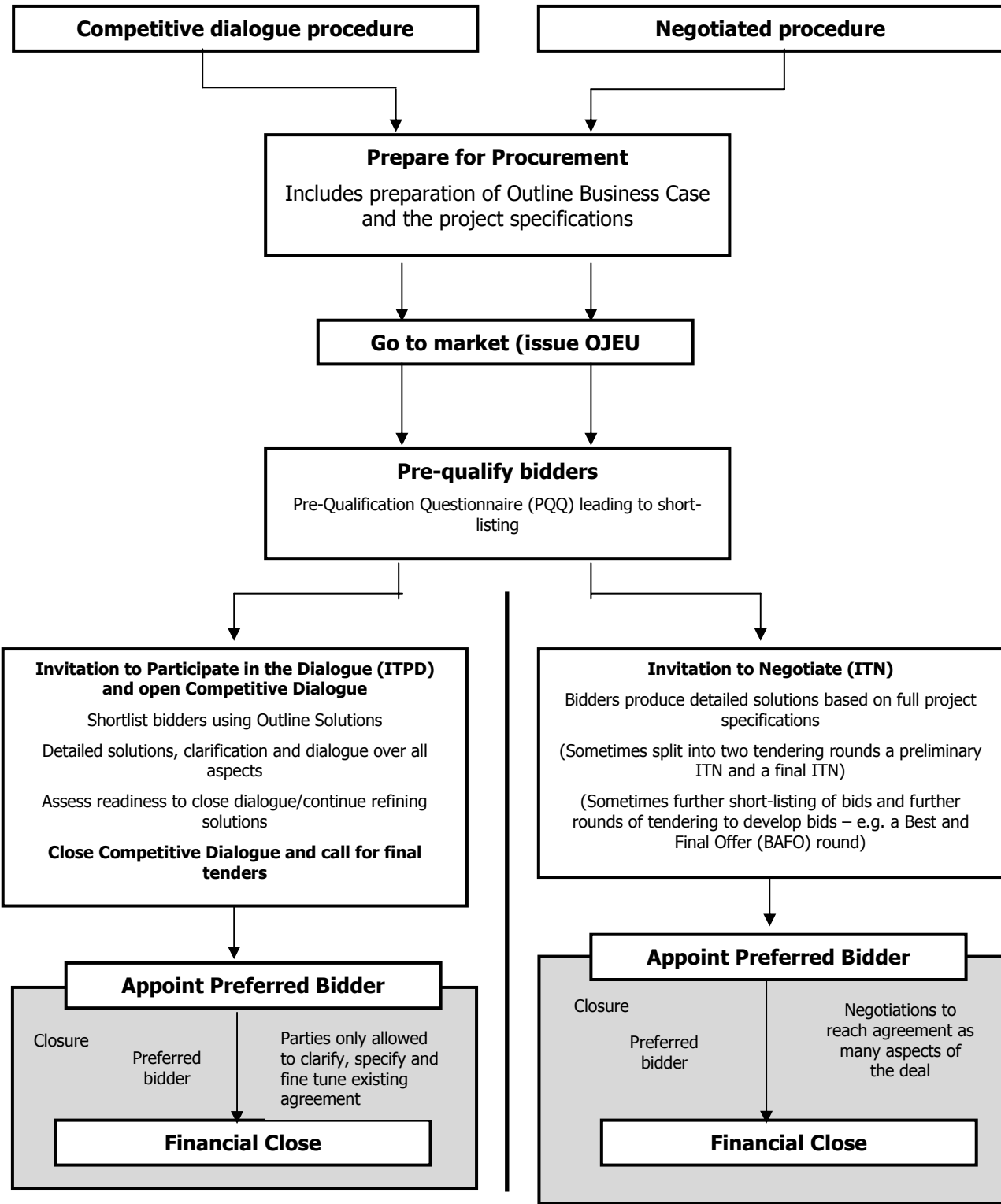
- i. CSFs provide a reliable monitoring mechanism for senior management level to identify key result areas;*
- ii. CSFs facilitate a structured, top down analysis or planning process.*

### *Weakness:*

- i. Difficult to use if there is a lack of competency to apply the method;*
- ii. Validity of the CSF is in question due to elements of bias during the development stage;*
- iii. CSF provides a snapshot perspective rather than a comprehensive overview of the environment.*

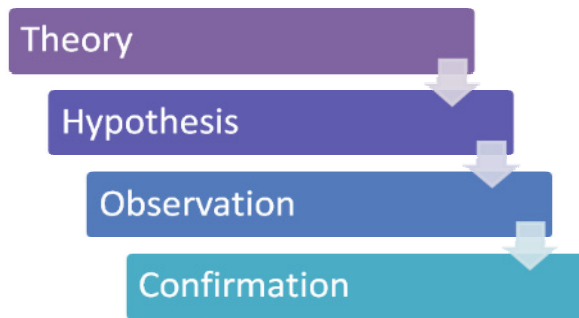
The application of CSFs in construction management and in PPP/PFI projects has been extensive. Zhang (2005) developed a suitable CSF package for PPP projects in general as these projects present a wide array of risks and complex relationship among the parties involved. Through the identification of CSFs, efficient allotting of available resources would contribute to the success of the project (Zhang, 2005). Li (2005) study on CSFs looked specifically into the PPP/PFI projects in the UK and identified 18 potential critical success factors that had an impact on the development of the PPP policy as well as future projects. Jefferies (2006) took a different approach by focusing on CSFs that affect the various participants in a PPP project environment through a case study. A study by Eaton (2006) narrowed the focus even further by concentrating on individuals in construction organisations involved in a PFI project. Eaton (2006) viewed creativity and innovation as improvements in PFI and categorised the factors that affect creativity and innovation into stimulants and

impediments which can be translated into CSFs. A deductive approach, see Figure 4.0, through an extensive literature review identifies the following four CSFs that are hypothesized to provide resilience to PFI, VfM, Risk Transfer, Affordability and Robustness (Figure 5.0).

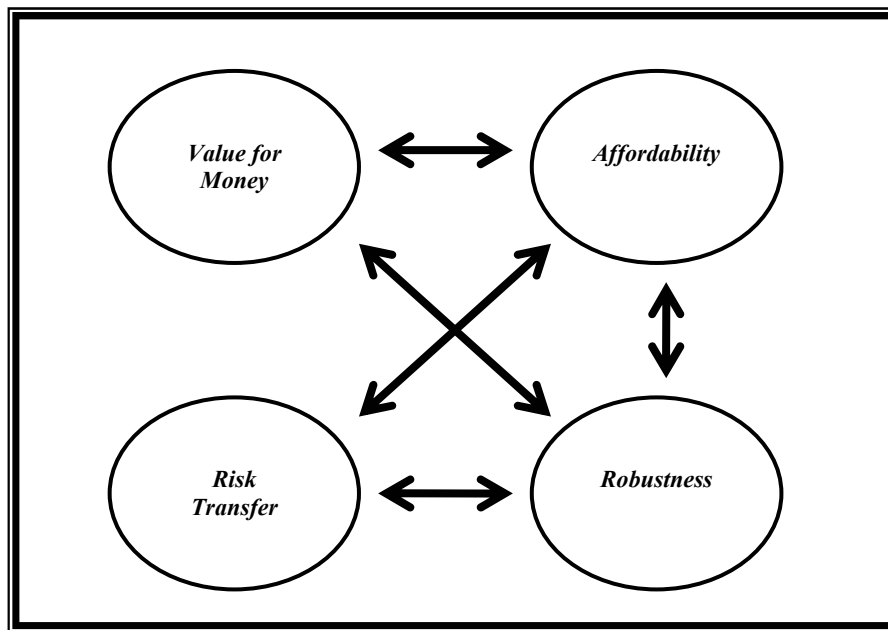


Source: Adapted from *National Audit Office, 2007*

**Figure 3.0** Comparisons between the Competitive Dialogue and Negotiated Procedure



**Figure 4.0** Deductive thinking approach



Source: Adopted by the authors from Akbiyikli, 2005

**Figure 5.0** Hypothesized Critical Success Factors

## 6.0 Value for Money, Risk Transfer, Affordability and Robustness

The termination of the Ryrie Rules in 1989 (Spackman, 2002) and the launch of PFI in 1992 (Payne, 1997) saw the term VfM being one of the core requirements for granting authorities to pursue the path of the off-balance sheet initiative. HM Treasury (2003:30) defines VfM as “*the optimum combination of whole-life cost and quality (or fitness for purpose) to meet the user requirement*”; in pursuing value for money, the Government plans to achieve:

- The evaluation of which procurement option to use is undertaken with no inherent preference for one option over another. There should be no dogmatism in this choice. Decisions should be made on the best evidence available;
- Value for money is not taken to be least cost. There is a need to ensure that quality standards are maintained, for example in the design of public infrastructure, and the long-term viability of the PFI contractor is assured;
- The commitment to value for money should not be at the expense of the items and conditions of employees transferred or subsequently employed by a PFI contractor; and
- A full evaluation of the costs and benefits on a whole-life basis is always undertaken, including an assessment of risk.

The approach of quantifying VfM in the UK has been through the Public Sector Comparator (PSC) (Sachs *et al.*, 2005). A study by the UK Government to measure the impact of VfM in

PFI projects was executed through Arthur Andersen and Enterprise LSE in 2000. The outcome of that study indicated an average savings of 17% being achieved in PFI procured projects evaluated comparative to traditional forms of procurement. Thus the VfM was seen to be driving economy, efficiency and effectiveness (Glendenning, 1988) but the reality was still being argued (Heald, 2002). The lack of a legal distinctive definition of VfM in the context of PFI points to the uncertainty of attaining VfM in PFI projects (Pitt *et al.*, 2006). Kirk and Wall (2001) claimed that VfM is achievable through a genuine risk transfer process from public to private sector, indicating the strong relationship between value for money and risk transfer. However the factors that leads to a successful PFI project is beyond monetary savings alone obtained from VfM (Asenova *et al.*, 2002), but rather a combination of factors and in this research it is perceived to be the four CSFs.

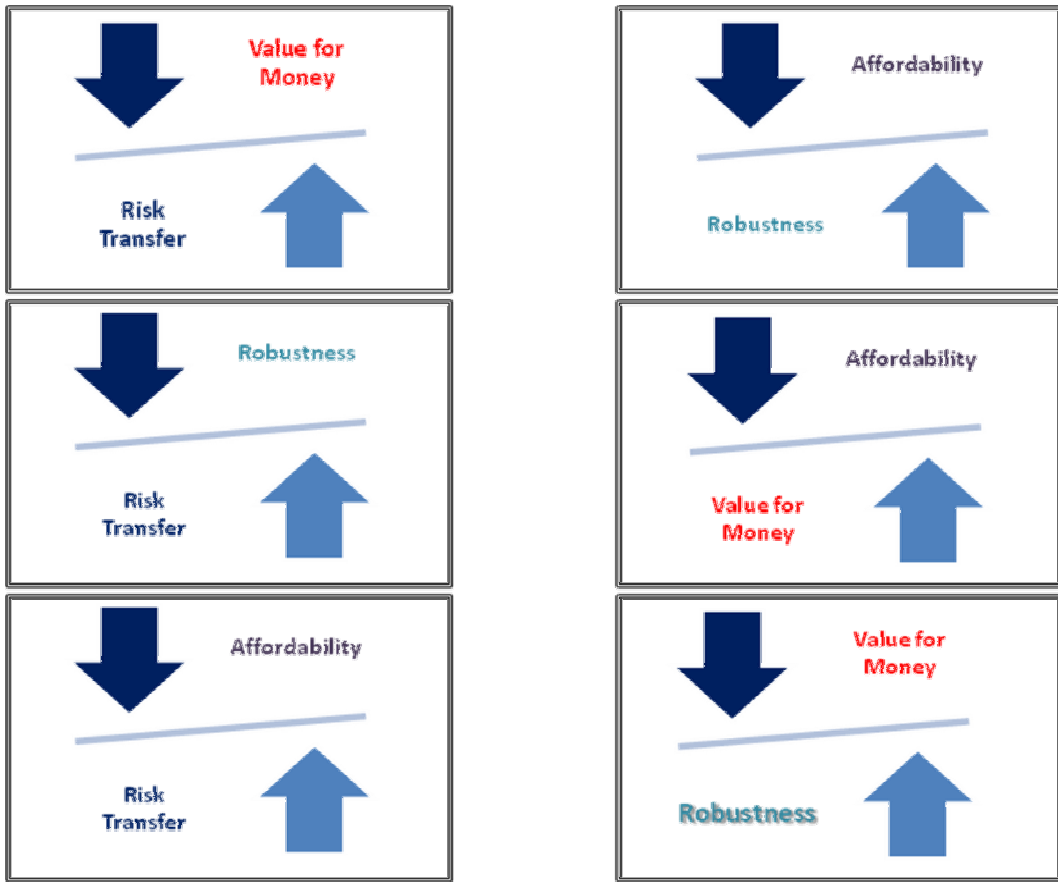
The operating mechanism of a PFI project transfers a series of responsibilities from the public to the private sector as part of the contract arrangement that allows the provision of public services by the private sector (Corner, 2006). The transfer of responsibilities is accompanied by the transfer of risk and the intensity of the risk varies throughout the whole-life cycle of a PFI project beginning from the project inception stage up to the provision of public services. However the process is more complex than just transferring 'lock stock and barrel' from public to private but rather allotting the risk to the party best able to manage it (Fox and Tott, 1999). Various studies have been conducted to look at different fields pertaining to PFI projects by various authors throughout the last three decades, among them Abednego and Ogunlana (2006); Akintoye *et al.* (1998); Akintoye and Chinyio (2005); Asenova and Beck (2003); Froud (2003); Gallimore *et al.* (1997); Ham and Koppenjan (2002); Sachs *et al.* (2007). The risk transfer exercise is coupled with elements of dynamic variety due to its nature, its intensity and occurrence within a PFI project environment. Failure to acknowledge the seriousness of risk transfer could lead to cost overrun, delay in time and level of quality.

The affordability factor is critically important in PFI projects due to the duration of such projects existing over a long term basis, 25-30 years and the responsibility of the granting authority as the purchaser of services to sustain the unitary payments to the service provider throughout the concession period (Fox and Tott, 1999; Shaoul, 2009). This reaffirms the definition by Grimsey and Graham (1997) that affordability is "the ability for the granting authority to budget for the projected unitary service fee required by the PFI project sponsor when taking into account central Government imposed financial constraints and the strategic and policy objectives of the granting authority". With the affordability factor behind the scenes, the need to satisfy the VfM factor is insufficient as it is the initial hurdle of two parts of the financial sustainability in a PFI project. Thus the evaluation of a PFI project environment should account for both VfM and affordability simultaneously to provide sustainability.

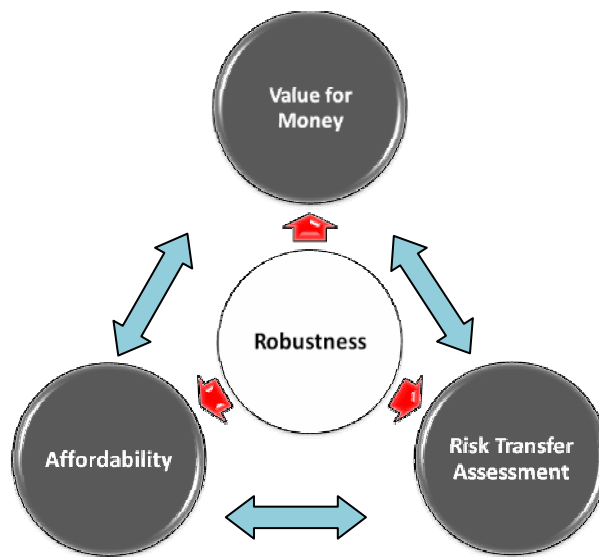
The engineering definition of robustness is described as the ability for a particular system to maintain its performance subjected to internal and external uncertainty parameters (Carlson and Doyle, 2002). Anderies *et al.* (2004) stressed that robustness accentuates the close relationship between cost benefits trade-offs and system design to cope with uncertainty. An allied concept to robustness is that projected by resilience which is defined as the persistence of relationships within a system and the measure of the ability of the system to absorb changes (Holling, 1973). Extending this approach of robustness into the PFI project environment it is vital to allow for viability over the long term concession period (Grimsey and Graham, 1997; Grimsey and Lewis, 2002). The presence of robustness from a whole-life cycle costing (WLCC) perspective within the context of a PFI project has been established in previous work (Akbiyikli, 2005). This can also be witnessed in the commitment by the UK Government towards ensuring that the VfM exercise achieves its objective through a rigorous assessment before going down the PFI path. Continuation in this direction is most certain as the confidence level among its users, providers and financiers increases over time.

The four CSFs, VfM, risk transfer, affordability and robustness are perceived to be the key contributing factors in the effective delivering of a PFI project. The significance and co-dependency of these four factors as a system will be pursued in this research to determine how they influence each other as well as their interdependency within a PFI project environment (Figures 6.0 & 7.0). The response and behaviour of these CSFs are governed by a series of drivers in the construction industry with regards to PPP and is defined through the SLEEPT methodology as discussed in the next section.





**Figure 6.0** CSFs Interdependenc



**Figure 7.0** CSFs Co-dependency

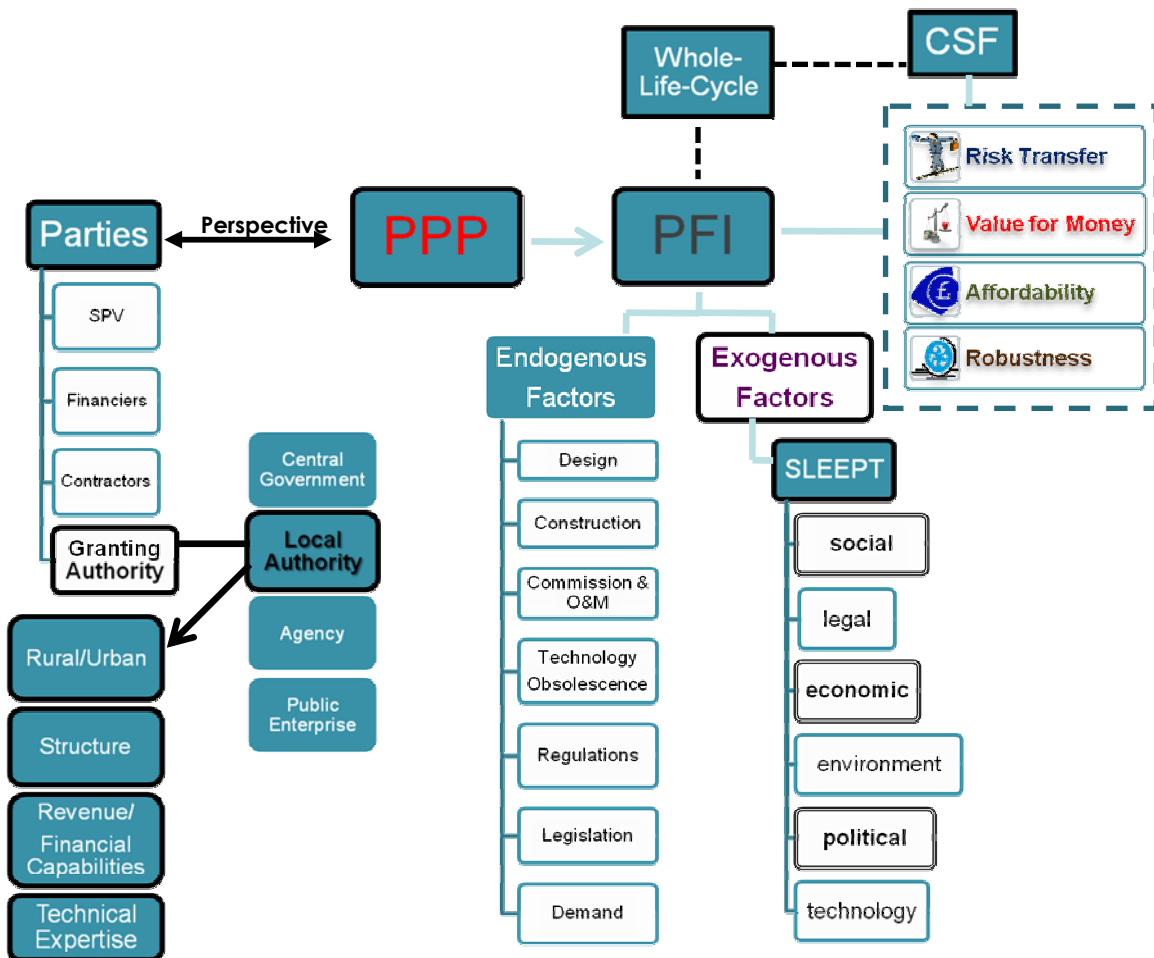
## 7.0 SLEEPT

The SLEEPT methodology has been established by The Centre for Risk Management Research (CRMR) at the University of Salford as a device for categorising 'drivers' of a process or object (Eaton *et al.*, 2006). The SLEEPT mnemonic is based on an isolation process of the activities concerned into six components both endogenous and exogenous of the unit of appraisal (Eaton and Akbiyikli, 2008). This research will only focus on the exogenous enablers of the unit of appraisal. The six components of SLEEPT are:

- Social
- Legal framework
- Economic
- Environmental
- Political framework
- Technological

S  
L  
E  
E  
P  
T

The role of SLEEPT in the research will contribute to the overall conceptual framework (Figure 8.0) of the PhD research towards determining the significance and co-dependency of robustness in a PFI project environment.



**Figure 8.0** Conceptual Framework

## 8.0 Conclusion

The conceptual framework of this PhD research is still at its preliminary stage. It is expected that as more literature knowledge is explored and obtained, the shape of the framework will be further refined to meet the aims and objectives set out at the beginning of this papers. It is envisioned that the framework would provide an insight in to how robust PFI is in order to provide sustainability to the initiative to stand the test of time. Further insights on how the key factors of Value for Money, Risk Transfer, Affordability and Robustness influenced by the SLEEPT drivers are able to cohabit and adapt to changes from the context of whole-life cycle are expected. The architecture required of a PFI to ensure its robustness from the granting authority's perspective is also explored to obtain valuable information that will assist granting authorities in general in their approach to utilising the PFI initiative for the provision of public services.

## References

- Abednego, M. P. & Ogunlana, S. O., (2006) *Good project governance for proper risk allocation in public-private partnerships in Indonesia*. International Journal of Project Management, 24 622-634.
- Akbiyikli, R., (2005) *The Holistic Realisation of PFI Road Project Objectives in the UK*. Unpublished PhD. School of Construction and Property Management. Salford, University of Salford.
- Akintoye, A. & Chinyio, E., (2005) *Private Finance Initiative in the healthcare sector: trends and risk assessment*. Engineering Construction & Architectural Management, 12 (6), 601-616.
- Akintoye, A., (2009) *PPPs for Physical Infrastructure in Developing Countries*. Akintoye, A. & Beck, M. (Eds.) Policy, Finance & Management for Public-Private Partnerships. Wiley-Blackwell.
- Akintoye, A., Taylor, C. & Fitzgerald, E., (1998) *Risk analysis and management of Private Finance Initiative projects*. Engineering Construction & Architectural Management, 5 (1), 9-21.
- Akintoye, A., Fitzgerald, E. & Hardcastle, C., (1999) *Risk management for local authorities' private finance initiative projects*. COBRA.
- Allen, G., (2003) *The Private Finance Initiative (PFI)*. Section, E. P. a. S. (Ed.). House of Commons Library.
- Anderies, J. M., Janssen, M. A. & Ostrom, E., (2004) *A Framework to Analyze the Robustness of Social-ecological Systems from an Institutional Perspective*. Ecology and Society, 9 (1), 18.
- Asenova, D. & Beck, M., (2003) *Scottish Local Authorities and the Procurement of Private Finance initiative Projects: A Pattern of Developing Risk Management Expertise?* Public Works Management Policy, 8 (11), 11-27.
- Asenova, D., Beck, M., Akintoye, A., Hardcastle, C. & Chinyio, E., (2002) *Partnership, Value for Money and Best Value in PFI Projects: Obstacles and Opportunities*. Public Policy and Administration, 17 (5), 5-19.
- Ball, R. & King, D., (2006) *The Private Finance Initiative in Local Government*. Economic Affairs. Blackwell Publishing Limited.
- Boynton, A. C. & Zmud, R. W., (1984) *An Assessment of Critical Success Factors*. Sloan Management Review, 25 (4), 17-27.
- Broadbent, J. & Laughlin, R., (2003) *Public private partnership: an introduction*. Accounting, Auditing & Accountability Journal, 16 (3), 332-341.
- Carlson, J. M. & Doyle, J., (2002) *Complexity and robustness*. Proceedings of the National Academy of Sciences of the United States of America, 99 (Suppl 1), 2538-2545.
- Corner, D., (2006) *The United Kingdom Private Finance Initiative: The Challenge of Allocating Risk*. OECD Journal on Budgeting, 5 (3).
- De Lemos, T., (2002) *Sustainable Competitive Advantage in PFI: A Systematic and Holistic Approach to Identify the CSFs in Risk Management in PFI, taking into account the Whole Life Cycle*. Unpublished PhD. Research Centre for the Built and Human Environment School of Construction & Property Management. Salford, University of Salford.
- De Lemos, T., Betts, M., Eaton, D. & Almeida, L. T. D., (2003) *The Nature of PFI*. The Journal of Structured and Project Finance, 9 (1), 59-74.
- Dick, W. & Akintoye, A., (1996) *Private finance initiative procurement method*. COBRA.

- Dixon, T., Pottinger, G. & Jordon, A., (2005) *Lessons from the private finance initiative in the UK: Benefits, problems and critical success factors*. Journal of Property Investment & Finance, 23 (5), 412-423.
- Dixon, T. J. & Pottinger, G., (2006) *Lessons from the real estate partnerships in the UK: Drivers, barriers and critical success factors*. Property Management, 24 (5), 479-495.
- Dorn, N. & Levi, M., (2008) *Do European procurement rules generate or prevent crime?* Journal of Financial Crime, 15 (3), 243-260.
- Eaton, D., (2005) *Quantifying quality: A report on PFI and the delivery of public services*. RICS.
- Eaton, D. & Akbiyikli, R., (2008) *The effect of risk transfer on PPPs: an evaluation based on a UK road project*. Davis, H. A. (Ed.) Infrastructure Finance Trends and Techniques. London, Euromoney Institutional Investor Plc.
- Eaton, D., Akbiyikli, R., Casensky, M., Sara, P. & Peterka, T., (2006) *Intra-Market Analysis of Facilities Management in UK Schools and Hospital PPP's: Lessons Learned*. McDermott, P. & Khalfan, M. M. A. (Eds.) Symposium on Sustainability and Value through Construction Procurement. Salford.
- Eaton, D., Akbiyikli, R. & Dickinson, M., (2006) *An evaluation of the stimulants and impediments to innovation within PFI/PPP projects*. Construction Innovation. Sage Publications, Ltd.
- Eaton, D., Akbiyikli, R., Lemos, T. D., Gunnigan, L., Kutanis, R. O., Casensky, M., Ladra, J. & Sawalhi, N. E., (2007) *An examination of the suitability of a UK PFI model within the Czech Republic, the Republic of Ireland, Palestine (Gaza-West Bank), Portugal and Turkey*. Construction Innovation: Information, Process, Management, 7 (1), 122-142.
- Ezulike, E. I., Perry, J. G. & Hawwash, K., (1997) *The barriers to entry into the PFI market*. Engineering Construction & Architectural Management, 4 (3), 179-193.
- Fourie, F. C. & Burger, P., (2000) *An Economic Analysis and Assessment of Public-Private Partnerships (PPPs)*. The South African Journal of Economics, 68 (4), 1-33.
- Fourie, F. C. & Burger, P., (2001) *Fiscal Implications of Public-Private Partnerships (PPPs)*. South African Journal of Economics, 69 (1), 147-168.
- Fox, J. & Tott, N., (1999) *The PFI Handbook*, Bristol, Jordon Publishing Limited.
- Froud, J., (2003) *The Private Finance Initiative: risk, uncertainty and the state*. Accounting, Organizations and Society, 28 567-589.
- Gallimore, P., Williams, W. & Woodward, D., (1997) *Perceptions of risk in the Private Finance Initiative*. Journal of Property Finance, 8 (2), 164-176.
- Ghobadian, A., Gallear, D., O'Regan, N. & Viney, H., (2004a) *PPP: The Instrument for Transforming the Public Services*.
- Ghobadian, A., Gallear, D., O'Regan, N. & Viney, H. (Eds.) *Public-Private Partnerships: Policy and Experiences*. Palgrave Macmillan.
- Ghobadian, A., Gallear, D., Viney, H. & O'Regan, N., (2004b) *Future of Public Private Partnership*.
- Glendenning, R., (1988) *The Concept of Value for Money*. International Journal of Public Sector Management, 1 (1), 42-50.
- Gosling, T., (2004) *3 Steps Forward, 2 Steps Back: Reforming PPPs*. New Economy, 11 (4), 229-235.
- Grimsey, D. & Graham, R., (1997) *PFI in the NHS*. Engineering Construction & Architectural Management, 4 (3), 215-231.
- Grimsey, D. & Lewis, M. K., (2002) *Evaluating the risks of public private partnerships for infrastructure projects*. International Journal of Project Management, 20 (2), 107-118.
- Grimsey, D. & Lewis, M. K., (2005) *Are Public Private Partnerships value for money?: Evaluating alternative approaches and comparing academic and practitioner views*. Accounting Forum, 29 (4), 345-378.
- Ham, H. v. & Koppenjan, J., (2002) *Building Public-Private Partnerships: Assessing and managing risks in port development*. Public Management Review, 4 (1), 593-616.
- Harding, A., Wilks-Heeg, S. & Hutchins, M., (2000) *Business, Government and the Business of Urban Government*. Urban Studies, 37 (5-6), 975-994.
- Heald, D., (2002) *Value for money tests and accounting treatment in PFI schemes*. Accounting, Auditing & Accountability Journal, 16 (3), 342-371.
- Hellowell, M. & Pollock, A. M., (2009) *Non-Profit Distribution: The Scottish Approach to Private Finance in Public Services*. Social policy & Society, 8 (3), 405-418.

- HM Treasury, (2000) *Public Private Partnerships: The Government Approach*. Treasury, H. (Ed.). HMSO.
- HM Treasury, (2003) *PFI: meeting the investment challenge*. Treasury, H. (Ed.). London, HMSO.
- HM Treasury, (2006) *PFI: strengthening long-term partnerships*. HM Treasury (Ed.). London, HMSO.
- HM Treasury, (2008) *Infrastructure procurement: delivering long-term value*. Treasury, H. (Ed.). London, HMSO.
- Hodge, G. A., (2004) *The risky business of public private partnerships*. Australian Journal of Public Administration. Blackwell Publishing Limited.
- Holling, C. S., (1973) *Resilience and Stability of Ecological Systems*. Annual Review of Ecology and Systematics, 4 1-23.
- Jefferies, M., (2006) *Critical succes factors of public private sector partnerships: A case study of the Sydney SuperDome*. Engineering Contruction & Architectural Management, 13 (5), 451-462.
- Klijn, E.-H. & Teisman, G. R., (2003) *Institutional and Strategic Barriers to Public Private Partnership: An Analysis of Dutch Cases*. Public Money & Management. Routledge.
- Kirk, R. J. & Wall, A. P., (2001) *Substance, Form and PFI Contracts*. Public Money & Management. Routledge.
- Li, B. & Akintoye, A., (2003) *An overview of public private-private partnership*. Akintola Akintoye, E. F., Cliff Hardcastle (Ed.) Public-Private Partnerships: Managing risks and opportunities. Blackwell Publishing.
- Li, B., Akintoye, A., Edwards, P. J. & Hardcastle, C., (2005) *Critical success factors for PPP/PFI projects in the UK construction industry*. Construction Management & Economics. Routledge.
- McKee, M., Edwards, N. & Atun, R., (2006) *Public-private partnerships for hospitals*. Bulletin of the World Health Organization, 84 890-896.
- NAO, (2007) *Improving the PFI tendering process*. Office, N. A. (Ed.). London, National Audit Office.
- Norman, G., (1990) *Life cycle costing*. Property Management, 8 (4), 344-356.
- Owen, G. & Merna, A., (1997) *The Private Finance Initiative*. Engineering Contruction & Architectural Management, 4 (3), 163-177.
- Pasquire, C. & Swaffield, L., (2002) *Life-cycle/Whole-life Costing*. Kelly, J., Morledge, R. & Wilkinson, S. (Eds.) Best Value in Construction. Blackwell Publishing.
- Payne, H., (1997) *Key legal issues in projects procured under the private finance initiative*. Engineering Contruction & Architectural Management, 4 (3), 195-202.
- Pitt, M., Collins, N. & Walls, A., (2006) *The private finance initiative and value for money*. Journal of Property Investment & Finance, 24 (4), 363-373.
- Pongsiri, N., (2002) *Regulation and public-private partnerships*. The International Journal of Public Sector Management, 15 (6), 487-495.
- Rockart, J. F., (1982) *The changing role of the Information Systems Executive: A Critical Success Factors Perspective*. Sloan Management Review, 24 (1), 3-13.
- Sachs, T., Elbing, C., Tiong, R. L. K. & Alfen, H. W., (2005) *Efficient Assessment of Value for Money (VFM) for selecting effective Public Private Partnerships (PPP) solutions - A comparative study of VFM Assessment for PPPs in Singapore and Germany*. Sidwell, A. C. (Ed.) The Queensland University of Technology Research Week International Conference. Brisbane Queensland University of Technology.
- Sachs, T., Tiong, R. & Wang, S. Q., (2007) *Analysis of political risks and opportunities in public private partnerships (PPP) in China and selected Asian countries*. 2, 1 (126-148).
- Shaoul, J.,(2009) *Using the Private Sector to Finance Capital Expenditure: The Financial Realities*. Akintoye, A. & Beck, M. (Eds.) Policy, Finance & Management for Public Private Partnerships. Chichester, Wiley-Blackwell.
- Spackman, M., (2002) *Public-private partnerships: lessons from the British approach*. Economic Systems, 26 (3), 283-301.
- Swaffield, L. M. & McDonald, A. M., (2008) *The contractor's use of life cycle costing on PFI projects*. Engineering Contruction & Architectural Management, 15 (2), 132-148.
- Zhang, X. Q., (2005) *Critical Success Factors for Public Private Partnerships in Infrastructure Development*. Journal of Construction Engineering & Management. American Society of Civil Engineers.

# Remediated Swing and the Celestial Jukebox: new 'friends' for jazz?

Tom Sykes, School of Music, Media and Performance, University of Salford

## Abstract

The almost hegemonic dominance of digital technology in contemporary Western life has affected music, particularly popular music, more than many other aspects of our culture. The coercing of a reluctant record industry into the transition has been widely discussed (particularly in the news media), but how has the impact of the internet affected the perception, dissemination and consumption of jazz? This paper looks at the possibilities for cross-disciplinary research in this area by drawing on new media and 'remediation' theory, the concept of the 'Celestial Jukebox', and community studies. It is hoped that this will uncover new paths for jazz studies, in a rapidly evolving digital culture.

Keywords: remediation, Celestial Jukebox, peer-to-peer, online communities

## Introduction

Information technology, and the way we use it, is an ingrained part of our culture and everyday activities. Analogous to the history of written communication, Timothy Taylor (2001) outlines the milestones in music history such as the invention of moveable type for music printing, and gramophone recording. He boldly describes digital technology as 'what may be the most fundamental change in the history of Western music since the invention of music notation in the ninth century' (2001: 3). The digitization of music has revolutionized the way it can be disseminated and stored, and the potential effects of this on the music industry in general are well documented.<sup>1</sup> But how is it affecting jazz? In the Jazz Services-commissioned report *The Value of Jazz in Britain*, what is described as 'a significant jazz presence on the internet' (Riley and Laing, 2006: 20) appears to be mainly one of a growing online CD market rather than a result of legal downloading. This is partly because 'some label owners believe that the lesser audio quality of MP3 compared to CD is off-putting to some jazz fans' (2006: 18). Whether this is a factor or not, the so-called 'long tail' effect of companies such as Amazon, whereby the sheer abundance of music available will create at least some sales for everything, is cited by Riley and Laing (2006: 18) as being, according to some in the industry, a positive thing. The 'long tail' theory has recently been questioned by Will Page of PRS for Music (the music royalties collection society), whose study of digital download sales found that millions of available tracks have failed to find a single buyer, although until the results are published it is not clear how many of these are in the category of 'jazz' (Foster 2008: 17).<sup>2</sup> This paper will consider the subject of jazz and the internet, using theoretical ideas borrowed from new media studies and debates about the changing state of the recording industry, presenting them as 'snapshots' of possibilities for jazz research.

## New media, old recordings

The first concept under consideration is that of 'remediation', based on a theory of Marshall McLuhan's (1964)<sup>3</sup> and which Jay Bolter and Richard Grusin define as 'the representation of one medium in another' (quoted in Auslander 1999: 7).<sup>4</sup> This is most evident when a new medium first appears, examples being early television replicating theatre (Auslander 1999: 13) and web sites resembling magazines (Lister et al 2003: 390). Conversely, an old medium can reinvent itself by incorporating a new one, such as CGI in cinema or the use of large screens showing continuous footage of a live concert for the audience distant from the stage (Auslander 1999: 24). In their early days, digitally recorded CDs were often sold alongside the

---

<sup>1</sup> See, for example, Biddle et al (2002), Dolfsma (2000) and Frith (2000).

<sup>2</sup> For more information on the 'long tail' theory see Anderson (2006).

<sup>3</sup> Specifically: 'A new medium is never an addition to an old one, nor does it leave the old one in peace. It never ceases to oppress the older media until it finds new shapes and positions for them' (McLuhan 1964: 158).

<sup>4</sup> Originally published in Bolter and Grusin (1996).

very same recordings on vinyl, thereby incorporating the new recording technology in the old playback format. Of course, recording itself could be considered a remediation of a live medium such as jazz, but with recorded music being a primary means of discovery, musical influence and consumption of jazz, it could be (and has been) argued that jazz and recording have had a symbiotic relationship throughout most of the music's history.<sup>5</sup> If this is the case, can the idea of remediation be then applied to jazz *recordings* in this world of *new media*?

An ontological discussion of what new media is (or are) is outside the scope of this paper, but Martin Lister et al (2003: 12) suggest a list of what is new about new media, which includes: new textual experiences; new relationships between users and media technologies; new experiences of the relationship between embodiment, identity and community; and new patterns of organisation and production. By these descriptors, any audio recording that is then converted to an MP3 file, so that it can be downloaded to an individual's playlist or shared among an online community would certainly fit within the 'new media' category, but to the extent to which it has actually been remediated is open to debate. On encountering the concept of remediation, I was initially unconvinced that simply converting audio that was already in digital form into MP3 format was actually remediating it. However, Taylor (2001: 3-4) points out the ease with which perfect copies of these files can be made and transmitted (particularly over the internet) in a *non-physical* way, and it is this latter characteristic that arguably distinguishes MP3s as remediations of CDs.

### **Pennies from heaven?**

Another new media concept integral to this study is that of the 'Celestial Jukebox'. The use of this term goes back at least as far as the Clinton administration's White Paper of 1995, and is defined by Patrick Burkart and Tom McCourt (2006: 1) as 'the various systems whereby any text, recording, or audiovisual artefact can be made available instantaneously ... to Internet appliances or home computers.'<sup>6</sup> The utopian dream of being able to access anything has largely been realized, but at a cost; examples are numerous, from unsubstantiated misinformation on Wikipedia to child pornography. As Burkart and McCourt (2006: 3-4) point out, the metaphorical implications of the term 'Celestial Jukebox' are of 'heavenly attributes, a gift from God (or, perhaps, the Big Four deities of the recording industry)', and the commercial potential of all this availability.<sup>7</sup> It is the operation of the Celestial Jukebox that may well determine the future of recorded jazz in the digital domain.

Central to the operation of the Celestial Jukebox is copyright. The recording industry has a long-established practice of using copyright law to protect its own interests and maintain control over the sale and distribution of its product, which before downloading was manifestly a *physical* product. Copyright law in America, as described by Jessica Litman (2006: 124), was originally designed as a 'bargain' between authors and the public, in which there was a balance between the legal 'fair use' the public could make of a work (which eventually included private copying, being difficult to prevent) and the limited period in which the author and publisher could exploit a work for profit, encouraging the creation of more new work. Copyright law has steadily been revised to *increase* the period under which a work retains copyright (an 'evolution in metaphors', as Litman puts it, 129), and the digital media industry in America has benefited from a significant extension in intellectual property rights as a result of the Digital Millennium Copyright Act of 1998. This act also disregarded fair use policy by criminalizing peer-to-peer computer file sharing, which coincided with the recent privatization and consequent growth in commercialization of the internet (Burkart and McCourt 2006: 7-9).

Much of the recent debate about the legality of file sharing has centred on what the content industry regards as 'piracy'; rather than simply being the illegal reproduction of works to be sold for profit, the industry now regards piracy to include any form of unauthorized copying for personal use or sharing with friends.<sup>8</sup> Record companies equate file sharing on peer-to-peer networks with piracy because, in Litman's words, 'if Johnny Teenager were to decide to share

---

<sup>5</sup> See, for example, Shipton (2001: 548-57), Katz (2004: 72-84) and Cuscuna (2005).

<sup>6</sup> The 'Celestial Jukebox' is referred to in Lehman (1995: 222) and discussed by Goldstein (2003: 184-5 and 187, first published 1996).

<sup>7</sup> The 'Big Four' refers to the four major record labels: Universal, Sony BMG, Time-Warner and EMI.

<sup>8</sup> See Litman (2006: 129) and Marshall (2004: 191) for music industry definitions of 'piracy'.

his unauthorized copy with two million of his closest friends, the *effect* on a record company would be pretty similar to the effect of some counterfeit CD factories creating two million CDs and selling them cheap' (2006: 129, emphasis in original). Hence the industry's highly publicized lawsuit against file sharing service Napster in 2000, where it was claimed that although Napster's users made no financial gain from the service, the fact that they could search for and share music files had led to decreased CD sales near American campuses where Napster was popular.<sup>9</sup> In its defence, Napster cited the 1992 Audio Home Recording Act, which allowed digital music transfer for personal, non-commercial use; the so-called Betamax case (1984), where the provider is not liable for technology that allows 'substantial non-infringing uses' (in this case transferring MP3s from hard disk to player and distributing non-copyrighted music); and even the First Amendment, the free-speech right to disseminate information, which in Napster's case was its indexing service (Burkart and McCourt 2006: 55-63). The BMG record company eventually tried to broker a business deal with Napster, which was ultimately unsuccessful, but the whole case illustrates the significance of legislation when studying new media and in shaping the Celestial Jukebox.

The 'alternative Celestial Jukebox' created by Napster may have been seen as being beneficial to CD sales if it could be shown that its users were sampling music like radio listeners and then buying CDs of the music they liked. Studies made at the time indicated that fewer than half of America's Napster users were young people aged 18 to 29, and that relatively few college students admitted to using the service to download music. However, this largely anecdotal evidence (discussed in Burkart and McCourt 2006: 56) was not obtained in a particularly scientific way, and the record industry took little notice of it. David Kusek and Gerd Leonhard's book *The Future of Music* (2005) decries the resistance of the record industry to change and its desperate protectionism, and offers some interesting proposals for new licensing of digital music streams, in a world of 'music like water' (2005: 3). Kusek and Leonhard are careful to make the distinction between the *music* business, which they believe has a healthy future, and the *record* industry, which they describe as 'terminally ill'. More importantly for the subject of this paper, they believe that 'the Internet, and digital networks in general, are starting to flip the niche genres from the bottom to the top' (2005: 7). This is perhaps a little optimistic, but the technology has certainly provided easier access to all styles of recorded music.<sup>10</sup>

### **The new communities**

Associated with file sharing networks is the idea of 'communities' in cyberspace. Nicholas Janowski (2006: 55) writes about the hopes of a resurrection of communities since the early days of radio, later, television, and in particular community radio and television in the 1970s. The academic study of this so-called 'community media' is well-established – even though, as Janowski (2006: 59) points out, the term 'community' is ill-defined sociologically – and continues within new media studies, with reference to virtual communities, by scholars such as Howard Rheingold and Steve Jones. Citing Jones (1995), Janowski (2006: 62) makes the point that, unlike their traditional counterparts, virtual communities are not usually geographically based, and that they can be described, in a way Allucquere Rosanne Stone suggested as long ago as 1991, as 'incontrovertibly social spaces in which people still meet face-to-face, but under new definitions of both "meet" and "face"'.<sup>11</sup> In their definition of online community, Jan Fernback and Brad Thompson (1995: n.p.) further nuance such a cyberspace relationship as 'forged through repeated contact within a specified boundary or place (e.g. a conference or chat line) that is symbolically delineated by topic of interest'. Jan Van Dijk (1998: 44-5) compares virtual communities with real-life (what he calls 'organic') communities, an example of their differences being that while an organic community tends to have a tight age group and several activities, a virtual community has loose affiliation and special activities (such as a specific musical interest). However, he concludes that virtual communities 'will not *replace* organic communities and ways of sociability, they will be *in addition to them*, build on them

---

<sup>9</sup> This is based on anecdotal evidence, discussed in Burkart and McCourt (2006).

<sup>10</sup> To date, little appears to have been written about the effect of digital technology on the success of niche music genres, other than suggestions made by authors such as Kusek and Leonhard (2005), Burkart and McCourt (2006) and Marcus (2003: 19), which are largely directed at the industry.

<sup>11</sup> Stone (1991: 85), quoted in Janowski (2006: 62).



and possibly strengthen them' (Van Dijk, 1998: 60, emphasis in original).<sup>12</sup>As peoples' discovery of music increasingly occurs via the internet, the study of online communities of fans (and musicians) will become increasingly important to music scholars.<sup>13</sup>

### **The death of the jazz album?**

So far, I have introduced these topics in the context in which they have mostly been written; that is, about performance, music, new media and contemporary culture in general. What are the possibilities for jazz? Jazz differs from pop music in that there has remained more of a market for CDs.<sup>14</sup> The matter of jazz record labels using the profits of back-catalogue re-releases to subsidize new recordings has been discussed by, among others, Michael Cuscuna (2005), Stuart Nicholson (2005) and Simon Frith (2007). Tony Whyton (2008: 158) writes about the reissue of 'classic' jazz recordings, which, along with 'bonus tracks and master editions feed the consumer desire for the "authentic" experience', at least for consumers within the historical, 'classic jazz' market. So what is happening to recorded jazz? Perhaps it is something like this: the much debated canon of jazz recordings is being exploited by the reissue industry (which caters for one type of jazz audience), while at the same time being (literally) deconstructed via individual track downloading, and then, arguably, reconstructed into new, personal canons by another type of jazz audience. As Michael Morse (2007: 167-8) points out, 'iPods and MP3 players eliminate even the minimal generic commitment of a CD, which presents the listener with an hour's worth of the "same" kind of music. The sole overt criterion for inclusion on an iPod playlist is "I like it", I want it there.' The Celestial Jukebox prospect for small jazz labels and individual musicians and bands is that there is more scope for niche marketing, but the lack of physical packaging and the individual track nature of online music may threaten the future of the album. If this leads to the return of single tracks marketed in the way 78s were, perhaps we will witness the history of recorded jazz coming, in a way, full circle.<sup>15</sup>

### **Conclusion**

Even from this brief survey it is evident that there are many lines of enquiry resulting from the encroachment of a twenty-first century digital culture on a twentieth century art form. Despite the optimism of those such as Kusek and Leonhard, who see an opportunity for 'minority' genres such as jazz to become popular as long as the record industry is prepared to change, there are others such as Gustavo Azenha (2006), who are more cautious: 'in reality the hierarchical priorities of major record labels have continued to limit the diversity of authorised music available online.'<sup>16</sup> Azenha also argues that, due to the political socio-economic contexts surrounding the use of information technology, decentralization of distribution does not necessarily lead to democratization of access, and that the record industry is taking advantage of its gate-keeping powers, which have in some ways been enhanced by technology.<sup>17</sup> An example of this is Digital Rights Management (DRM), which ideally, Adam Marcus (2003: 2) suggests, 'would be invisible to the licensed owner of content, but would make it impossible for others to access the same content.' However, Marcus disagrees with over-zealous DRM and

---

<sup>12</sup> The question of the extent to which this has occurred would be a possible avenue for research.

<sup>13</sup> A survey of musicians was carried out by the Pew Internet and American Life Project (Madden 2004) (which was, by its own admission, 'self-selecting and non-projectable onto the general population of musicians', ii), and among its findings were: that musicians 'use the internet to gain inspiration, build community with fans and fellow artists, and pursue new commercial activity' (ii); that 'the vast majority do not see online file-sharing as a big threat to creative industries' (ii); and that 'like most internet users, online artists are also active consumers of media content online' (vii). Certainly, the widespread use of networking sites has led to increased opportunity for cheap publicity among musicians.

<sup>14</sup> The British jazz situation, although no longer up to date, is summarised in Riley and Laing (2006). The existence of a current market for jazz CDs has recently (May 2009) been confirmed to me anecdotally by industry insiders including Christine Allen (JazzCDs and Basho Records) and Michael Cassidy (I Think Music), both of whom sell jazz CDs via their web sites.

<sup>15</sup> 78 r.p.m. records were the pop singles of their time, which coincided with the period when jazz was at its most popular, between the 1920s and 1940s (see Shipton 2001: 554-7).

<sup>16</sup> Azenha (2006: n.p., para. 108).

<sup>17</sup> Azenha (2006: n.p., para. 4-5).

advocates a subscription model similar to that of Kusek and Leonhard. Since these authors' work was published, there have been some significant developments in the industry, such as the agreement by Apple (which has a large share of the digital music market) to allow DRM-free downloads from its iTunes store, thereby removing the restriction of being able to play them only on iTunes software or an iPod (unlike many of Apple's competitors, which have already removed DRM). This follows the recent introduction of downloadable, DRM-free music by one of the largest online sellers of CDs, Amazon.

There are also current political developments that will directly affect producers, distributors and consumers of digital media in Britain. In its interim report on *Digital Britain* (Carter 2009: 7-13), the Department for Culture, Media and Sport presents a list of actions which include: the possible establishment of a Rights Agency to 'enable copyright-support solutions that work for both consumers and content creators' (including fair use policy); consultation on a legislative approach to peer-to-peer file sharing; and the availability of (and participation in) digital media including a National Media Literacy Plan. The government's intention still appears to be one of taking legal action against illegal file sharers rather than introducing a subscription-based licensing system, so Marcus and Kusek and Leonhard may have to wait a little longer for their ideal Celestial Jukebox, at least in Britain.

Where does this leave jazz scholars? If jazz is increasingly treated as any other type of popular music by the music industry and consumers, then perhaps, as Frith (2007: 22) suggests, popular music scholars should include more jazz in their studies. By the same token, jazz scholarship should engage with some of the issues, such as those surrounding digital technology and the recording industry, being tackled by popular music and new media academics. In doing so, jazz scholars' understanding of new media and the analysis of its relationship to jazz should lead to a better understanding of jazz audiences and their modes of consumption. This will involve reference to theoretical models from disciplines that may be new to many jazz scholars (myself included), but by doing so it will offer new paths for jazz research and discourse, and conceivably also bring jazz more seriously into popular music studies. After all – who knows? – the Celestial Jukebox may yet bring remediated Swing back into the popular music mainstream.<sup>18</sup>

## References

- Anderson, Chris. 2006. *The Long Tail: How Endless Choice is Creating Unlimited Demand*. London: Random House Business.
- Auslander, Philip. 1999. *Liveness: Performance in a mediatized culture*. London: Routledge.
- Azenha, Gustavo S. 2006. 'The Internet and the decentralisation of the popular music industry: critical reflections on technology, concentration and diversification'. In *Radical Musicology*, [www.radical-musicology.org.uk/2006/Azenha.htm](http://www.radical-musicology.org.uk/2006/Azenha.htm) (accessed 19/2/09).
- Biddle, Peter, Paul England, Marcus Peinado, and Bryan Willman. 2001. 'The Darknet and the Future of Content Distribution'. Paper, ACM Workshop on Digital Rights Management, November 18, 2002. Originally published at [www.crypto.stanford.edu/DRM2002/prog.html](http://www.crypto.stanford.edu/DRM2002/prog.html) (accessed 16/2/09).
- Bolter, Jay David and Richard Grusin. 1996. 'Remediation'. In *Configurations*, 3: 311-58.
- Burkart, Patrick and Tom McCourt. 2006. *Digital Music Wars: Ownership and Control of the Celestial Jukebox*. Lanham: Rowman & Littlefield.
- Carter, Stephen. 2009. *Digital Britain: The Interim Report*. Norwich: The Stationery Office.
- Cuscuna, Michael. 2005. 'Strictly on the Record: the art of jazz and the recording industry'. In *The Source: Challenging Jazz Criticism*, 2: 63-70.
- Dolfsma, Wilfred. 2000. 'How Will the Music Industry Weather the Globalization Storm?' *First Monday*, [www.firstmonday.org/issues/special10\\_7](http://www.firstmonday.org/issues/special10_7) (accessed 15/2/09).
- Fernback, Jan and Brad Thompson. 1995. 'Virtual communities: abort, retry, failure?' Unpublished paper. Available online: [www.well.com/user/hlr/texts/Vccivil](http://www.well.com/user/hlr/texts/Vccivil) (accessed 2/3/09).

---

<sup>18</sup> For the benefit of those not familiar with jazz history, the so-called 'Swing era' of the 1930s was the last time a jazz style was considered the mainstream popular music of the day.

- Foster, Patrick. 2008. 'Sting in the tail for online sellers as 10m music tracks spin unloved through cyberspace'. In *The Times*, 22 December: 17.
- Frith, Simon. 2000. 'Music industry research: Where now? Where next? Notes from Britain'. *Popular Music*, 19/3: 387-93.
- . 2007. 'Is Jazz Popular Music?' In *Jazz Research Journal*, 1.1: 7-23.
- Goldstein, Paul. 2003. *Copyright's Highway: From Gutenberg to the Celestial Jukebox*. Stanford: Stanford University Press.
- Jankowski, Nicholas W. 2006. 'Creating Community with Media: History, Theories and Scientific Investigations'. In Leah A. Lievrouw and Sonia Livingstone (eds.), *The Handbook of New Media: Social Shaping and Social Consequences of ICTs*. London: Sage, 55-74.
- Jones, Steve. 1995. 'Understanding community in the information age'. In Steve Jones (ed.) *Cybersociety: Computer-mediated Communications and Community*. Thousand Oaks: Sage, 10-34.
- Katz, Mark. 2004. *Capturing Sound: How technology has changed music*. London: University of California Press.
- Kusek, David and Gerd Leonhard. 2005. *The Future of Music: manifesto for the digital music revolution*. Boston: Berklee Press.
- Lehman, Bruce A. 1995. *Intellectual Property and the National Information Structure: The Report of the Working Group on Intellectual Property Rights*. Washington DC: US Patent and Trademark Office.
- Lister, Martin, Jon Dovey, Seth Giddings, Iain Grant and Kieran Kelly. 2003. *New Media: A critical introduction*. Abingdon: Routledge.
- Litman, Jessica. 2006 [2001]. 'Choosing metaphors'. In Robert Hassan and Julian Thomas (eds.) *The New Media Theory Reader*. Maidenhead: Open University Press, 123-32.
- Madden, Mary. 2004. 'Artists, Musicians and the Internet'. *Pew Internet & American Life Project*, [www.pewinternet.org/](http://www.pewinternet.org/) (accessed 17/2/09).
- Marcus, Adam David. 2003. 'The Celestial Jukebox Revisited: Best Practices and Copyright Law Revisions for Subscription-Based Online Music Services'. Unpublished paper. Available online: [http://tprc.org/papers/2003/220/TPRC\\_paper-Adam\\_marcus.pdf](http://tprc.org/papers/2003/220/TPRC_paper-Adam_marcus.pdf) (accessed 15/2/09).
- Marshall, Lee. 2004. 'Infringers'. In Simon Frith and Lee Marshall (eds.) *Music and Copyright* (2<sup>nd</sup> edn.). Edinburgh: Edinburgh University Press.
- McLuhan, Marshall. 1964. *Understanding Media: The Extensions of Man* (2<sup>nd</sup> edn.). New York: New American Library.
- Morse, Michael W. 2007. 'Musical genre distinction and the uniculture: a reply to Simon Frith's "Is Jazz Popular Music?"'. In *Jazz Research Journal*, 1.2: 153-72.
- Nicholson, Stuart. 2005. *Is Jazz Dead? (Or has it moved to a new address)*. Abingdon: Routledge.
- Riley, Mykaell and Dave Laing. 2006. *The Value of Jazz in Britain*. London: University of Westminster.
- Shipton, Alyn. 2001. *A New History of Jazz*. London: Continuum.
- Stone, Allucquere Rosanne. 1991. 'Will the real body please stand up: boundary stories about virtual cultures'. In M. Benedikt (ed.) *Cyberspace: First Steps*. Cambridge, MA: MIT Press, 81-118.
- Taylor, Timothy D. 2001. *Strange Sounds: Music, Technology & Culture*. London: Routledge.
- Van Dijk, Jan. 1998. 'The Reality of Virtual Communities'. In *Trends in Communication*, 1 (1): 39-63.
- Whyton, Tony. 2008. 'Acting on Impulse!: Recordings and the Reification of Jazz.' In Mine Dogantan-Dack (ed.), *Recorded Music: Philosophical and Critical Reflections*. London: Middlesex University Press, 155-71.

# Using Virtual Learning Environment: Supporting Teaching and Learning, and Bridging the Digital Divide in Rural Communities in Ghana

**Najim Ussiph**, *School of Computing, Science and Engineering, University of Salford*

## **Abstract**

This project intends to use a rural community in Ghana, a typical developing economy in West Africa, to explore how ICTs can be used to tackle the problem of a lack of educational infrastructure and qualified teachers in rural community schools. These rural communities are often characterized by a lack of access roads or limited accessibility, unreliable or lack of power supply, low population and limited income.

In this research project, we intend to apply current advances in ICT to implement and extend wireless internet connectivity to the rural community under research. A model, in the form of a resource center for the research, is proposed in this paper. The resource center, which will also double as community e-center (CeC), will be based on a designated school premise. Regular e-learning packages or programs will be developed based on the curriculum needs of the school and external resource persons and specialist teachers will be used to deliver teaching over the net. A local area network with bus configuration will be installed at the e-center to network about 15 to 20 desktop PCs. Each PC will be connected to the Internet with the available bandwidth for high speed access using a switching hub. Other supporting hardware such as LCD projectors, scanners and printers will also be in use. The project plans to tackle the problem of power supply using diesel engine power generators. The e-center will serve the community by providing internet access to residents as well as evening educational programs on computer literacy and advantages of the internet.

The duration of the project will be one academic term and its impact will be assessed with respect to (i) students' motivation / interest, (ii) whether or not there has been improvement in academic performances, (iii) motivation of the teaching staff, (iv) work output or productivity of the teaching staff, (v) impact on school administration, (vi) ICT awareness of the community, (vii) benefit of the project to the community, and (viii) overall socio-economic impact on the community.

## **Introduction**

Under-development has been a major characteristic of the so called Third World countries in the last millennium and it is likely to continue well into the current millennium. Another dimension that is fast emerging in addition to already established facets of under-development of the third world countries is the *digital divide*.

The digital divide has been defined in several ways. It was described as the gap that exists in the opportunities to access advanced information and communication technologies between geographic areas or by individuals at different socio-economic levels (OECD, CICCIP, 2001). Harris (2004) expressed it as the uneven global distribution of access to the Internet that separates individuals who are able to access computers and the Internet from those who have no opportunity to do so and George Sciadas simply termed it as the gap between ICT "haves" and the "have-not's" (Sciadas, 2004). Broadly, the digital divide may be described as the gap between access and lack of access to information infrastructure as well as the basic information needs for day-to-day socio-economic activities among communities, groups or nations.

The advancement in technology in modern times, particularly in the area of Information and Communication Technology (ICT) and its permeation into all aspect of human life has brought about a new era often referred to as the Information Age. In this era of the Information Age, ICT services have come to be recognized as an important vehicle for the socio-economic development of any nation. The attainment of a higher level of development hinges on rapid progress of this technology as it opens several new opportunities for growth. ICT today has permeated every aspect of our lives, government and policy formulation from agriculture to national security down to welfare. Access to this information infrastructure should be no

privilege, it needs to be seen as a right in the much envisioned information society – a people-centered, inclusive and development-oriented information society, where everyone can create, access, utilize and share information and knowledge, enabling individuals, communities and peoples to achieve full potential in promoting sustainable development and improving their quality of life (WSIS, 2003). This concern could not have been better expressed than it was by the former UN Secretary-General, Kofi Annan at the WSIS 2003 summit in Geneva. He commented that “the new information and communications technologies are among the driving forces of globalization. They are bringing people together, and bringing decision makers unprecedented new tools for development. At the same time, however, the gap between information ‘haves’ and ‘have-nots’ is widening, and there is a real danger that the world’s poor will be excluded from the emerging knowledge-based global economy”. It was noted at the International Telecommunication Union (ITU) and the Korea Agency for Digital Opportunity and Promotion (KADO) Symposium On Building Digital Bridges (ITU/KADO, 2004: 3) that

Wireless technologies used for entertainment and convenience in developed world could have much more profound effects in developing world. While checking emails in a distant city from the back of a taxi surely makes business travel more effective in richer economies, being able to send and receive messages at all in a remote, rural community could produce much greater improvement in social welfare for the same infrastructure investment.

Therefore bringing the power and services of ICT to the rural communities that form up to 60 percent of most developing countries will in no small measure enhance economic development of such communities and the country as a whole. Staggering statistics (Harris, 2004) serve to highlight the alarming differences between those at both ends of the digital divide:

- All of the developing countries of the world own a mere four percent of the world’s computers.
- 75 percent of the world’s 700 million telephone sets can be found in the nine richest countries.
- There are more web hosts in New York than in continental Africa; there are more in Finland than in Latin America and the Caribbean combined.
- There were only 6.3 million Internet subscribers on the entire African continent in September 2002 compared with 34.3 million in the UK. (Nua Internet)

Table 1 shows the gap in Internet access between the industrialized and developing worlds. More than 85 percent of the world’s Internet users are in developed countries, which account for only about 22 percent of the world’s population.

**Table 1** Online Users (in millions) as of September 2002 (Harris, 2003: 9)

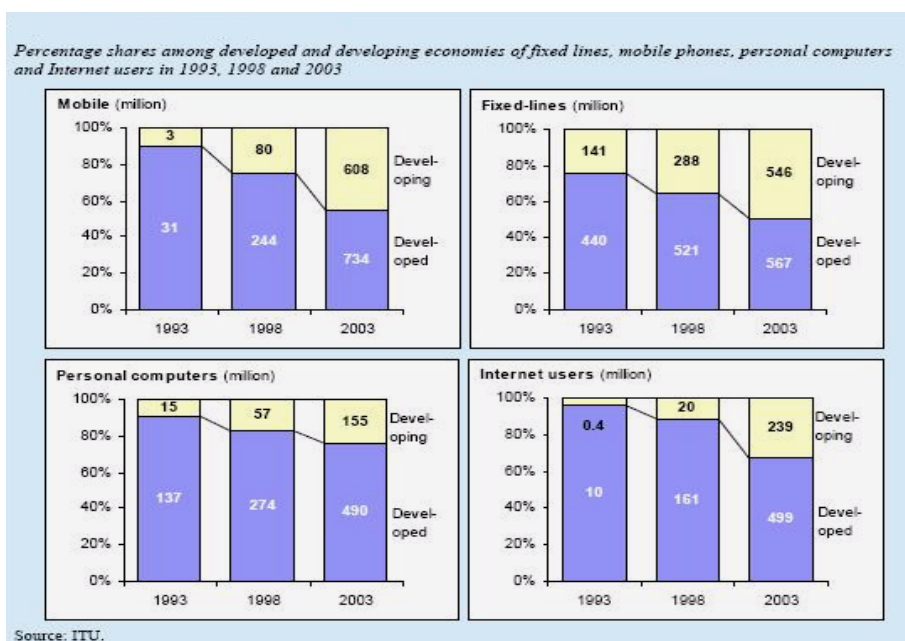
Africa	<b>6.31</b>
Asia/Pacific	<b>187.24</b>
Europe	<b>190.91</b>
Middle East	<b>5.12</b>
Canada & USA	<b>182.67</b>
Latin America	<b>33.35</b>
<b>World Total</b>	<b>605.60</b>

The good news here is that the developing countries have the opportunity of taking advantage of emerging information and communication technologies in the developed world and adapting them to provide ICT access and services in rural areas of developing countries at minimal cost. In this pursuit, we must not be oblivious of the fact that education, knowledge, information and communication are core to achieving this goal.

Even though statistics show that there has been improvement in the level of access to telecommunication facilities in the developing countries in the past five years, (see Figure 1) this improvement centered mainly in the urban areas. The rural communities are still in total darkness as far as ICT is concerned.

As ICTs began to penetrate our lives and the benefits, actual or potential, associated with their usage started to be understood, the undesirability of leaving behind substantial population masses surfaced as a key challenge of our times (Sciadas, 2004).

The frustration and challenge in trying to bridge the digital divide in a country or community however, is that the problem of the digital divide continues to exist as a result of rapidly advancing and emerging new technologies. Originally, the digital divide referred to access and lack of access to Internet connection over a dial-up connection at 56 kbit/s. Now, just a few years later in the developed economies the digital divide problem has shifted to unequal access to fast broadband access to the Internet (at a speed of 100 Mbit/s or more) and the applications higher-speed connection can provide (ITU/KADO, 2004). Thus, forming what looks like a never-ending catch-up for developing economies (Figure 2).

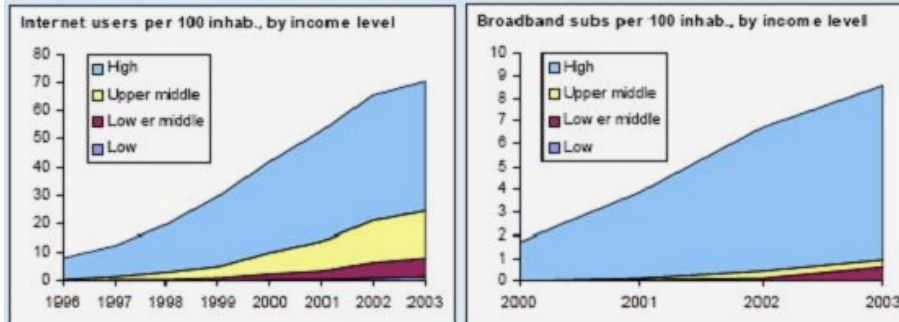


**Figure 1** The Shrinking Digital Access

*Just as developing economies start making progress bridging the digital divide in one ICT, new technologies appear and the cycle repeats.*

In mid 1990's the Internet was a new communication medium whose benefits were touted as the ultimate information tool for users. The Internet was the technology that was supposed to open libraries worth of materials to users at the click of a mouse. Some of the initial excited soon gave way to new fears about how people in developing economies would be able to make use of this new, incredible information tool. This was the emergence of the "digital divide" and policy makers went to great lengths to figure out how to prevent new classes of "haves" and "have-nots" in their economies and between nations.

However, just the numbers are starting to look more promising throughout the developing world for basic Internet access (left figure), a new technology, broadband, has appeared in markets around the world and is allowing users much "better" access than simple dial-up connections (right figure).



Source: ITU World Telecommunication Indicators Database

**Figure 2** The never-ending catch-up cycle for developing economies

We may then need to redefine digital divide to mean better access to ICTs for some people more than others. Just as the information revolution comes with the danger of marginalization within and across communities and nations, it has the potential to dramatically raise socio-economic status of a community and nation. Contemporary Information and Communication Technology (ICT) can now be used to integrate rural and poor urban communities into economic life, providing access to better education, health care, and raising income thereby improving living condition.

### Project Objective

This research project intends to explore and expose how e-learning (virtual learning environment) can be used to tackle the problem of a lack of educational infrastructure and qualified teaching personnel in rural community schools and also to bridge the digital divide in the rural communities. The research is a typical community informatics (Gurstein, 2000) research that seeks to provide new knowledge and insight into the means to enable and empower rural communities through the use of information and communication technologies (ICTs) and to move them beyond traditional barriers and impediments to find a new and more equitable role in the Information Society

### Research Environment

This research intends to use a selected rural community in Ghana, a developing country in the Sub-Sahara Africa as the research environment. Ghana is relatively a small country with a total area of 238,537 sq. km (92,100 sq. mi.). Land surface area of Ghana is 230,940 sq km and water surface area is 8,520 sq km. (GhabaWeb, 9/2/2008). Ghana has a population of 18.8 million (2000 Census) with population growth rate of 1.25 percent. The population estimate for 2005 was 21,029,853. The population density according to the 2000 Census was 78.9 persons/sq. km with average household size of 5.1. Urban population (1990) was 33 percent. Human Development Index (HDI) of Ghana is 0.562 (2002).

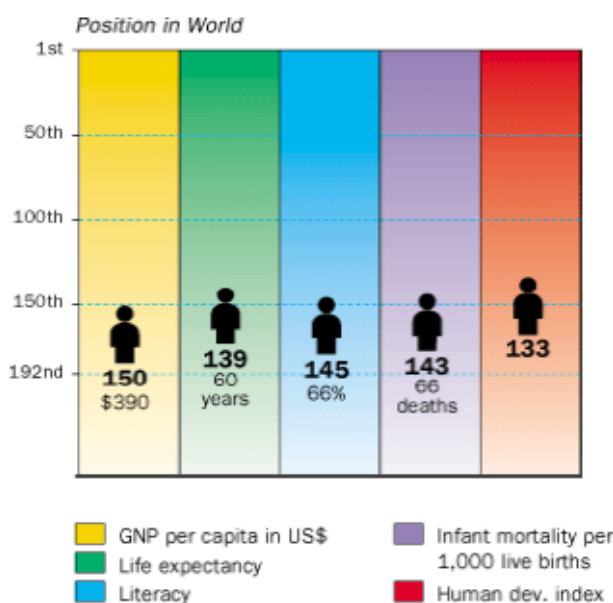
The alarming revelation of these statistics is the small size of the urban population. One can imagine the degree of digital exclusion for a country with 67 percent of the population living in the rural communities. The divide appears even more vivid when one considers the literacy rate in the country. The literacy rate is 74.8% (Male – 82.7%, Female – 67.1%). However, the 2001 National Population and Housing Census showed that "43.4 per cent of those who are

three years old or more have never been to school and 49.9 per cent of the adult population of 15 years or more are totally illiterate." (GhanaWeb, 11/03/08) The overall literacy rate therefore implies extremely low literacy rate in the rural areas.

Ghana has 12,130 primary schools, 5,450 junior secondary schools, 503 senior secondary schools, 38 teacher training colleges, 10 polytechnic institutions, 26 (6 public and 20 private) universities and one major degree awarding institute serving a learning population of 17 million (GhanaWeb, 11/03/08).

### E-Readiness of Ghana

Ghana has witnessed a steady growth in the provision of ICT infrastructure in the past 5 years. The World Economic Forum tracked and ranked Ghana at the 64<sup>th</sup> position in its 2004/5 report. The Report used the Networked Readiness Index – a measure of the degree of preparation of a nation to participate in and benefit from ICT developments. The country had been ranked 74 in 2003/4 (MoC Press Brief, Feb., 2006).



**Figure 3.** Ghana's ranking in the world (WEF, 2005)

Ghana's teledensity stands at 15 percent as at the end of 2005 implying that 15 out of every 100 people in Ghana have telephones. It was 1 percent in 2000. The breakdown of telephone usage in the country is as shown in Table 2 below. The government of Ghana is pursuing a program to introduce broadband to support the deployment and application of ICT country-wide. Particular attention is being given to the development of the Internet and to lower the cost of broadband infrastructure in Ghana.

The Ghana Internet Service Providers Association (GISPA) has established the Ghana Internet Exchange Point, located at the Kofi Annan Center of Excellence In ICT (KACE). There have been increases in Internet subscriptions and points-of-presence have been established in several parts of the country.

**Table 2** Distribution of Telephone Lines

Type	Number
Fixed Land lines	<b>331,000</b>
Payphones	<b>11,000</b>
Mobile Cellular lines	<b>2,655,000</b>
<b>Total</b>	<b>2,997,037</b>



Ghana Telecom (GT) – the only public and the main telecommunication service provider in the country, which has recently been taken over by the British giant, Vodafone – has begun to roll out broadband service in the capital, Accra, the nearby industrial city, Tema and in Kumasi. Corporate and urban customers in these cities can now have the GT Broadband service which is mainly provided on demand. From a figure of 1024 broadband lines in 2004, GT in 2005 delivered 8,600 broadband connectivity to businesses and homes to facilitate quick delivery of Internet based products and services. It was projected that the figure will rise to 19,400 in 2006 (MoC Press Brief, Feb., 2006).

The government with the support of The World Bank is also embarking on an e-Government program. The program is to develop a system that will use web-based Internet applications like web services in combination with other information processing technologies such as centralized database systems for the country to improve administrative services delivery and productivity and provide better coordination and interaction between government, the public and the private sector.

In the area of human capacity building, virtually all the universities and the polytechnic institutions in Ghana today offer courses in ICT. There is also the modern and highly equipped Advance Information Technology Institute (AITI), known as Ghana-India Kofi Annan Center of Excellence in ICT (KACE). This institution came to be as a result of collaboration between the governments of India and Ghana, initiated by the former UN Secretary-General, Kofi Annan. The government of India supported the project with hardware and technical assistance and gave Ghana its first supercomputer, the famous Indian-built Param Padma. The center offers Advanced Diploma courses in IT and provides training in capacity building for governmental agencies. It has also been made a Cisco Academy recently, the only one in the sub-region that train network professionals in Cisco network certification courses.

There are also other foreign private training institutions like NIIT and IPMC that offer courses leading to Diploma award and other vendor certification courses.

Despite all the above-mentioned strides to achieve the objectives of ICT4D and Vision 2010 in making Ghana a middle income country and set the path for becoming an information and knowledge - based economy, I can confidently assert that most of the rural communities that form about 67 percent of the Ghanaian population are still wallowing in near total digital exclusion.

The population distribution of Ghana as outlined above, the illiteracy rate and the distribution of the ICT infrastructure of the country reveal the magnitude of the digital divide in the country which thus call for a very comprehensive and vigorous agenda to increase literacy rate in the rural communities. The most viable means to pursue this agenda is to take advantage of the emerging information and communication technologies and their applications and particularly the Internet in education.

The application with perhaps the greatest influence on education and society is e-learning or online learning. E-Learning makes practicable a means of asynchronous collaborative learning, which until recently seemed an oxymoron in educational spheres (Garrison and Anderson, 2003).

Communication is at the core of knowledge transfer and any learning activity. The compelling advantage of e-learning is that it supports both synchronous and asynchronous communication in multimedia format (text, audio and video) and it can be independent of both time and distance.

Even advanced nations like the US could not resist the influence of e-learning and power of the Internet on its education. In 2001, the US Web-based Education Commission (on-line) had reported:

The question is no longer if the Internet can be used to transform learning in new and powerful ways. The Commission has found out that it can. The Web-based Education Commission calls for the new Congress to embrace an 'e-learning' agenda as a centerpiece of our country's federal education policy.

Also, the report of the Advisory Committee on Online Learning in Canada in 2001, the committee expressed that:

In a global society based on expanding knowledge, Canada's health as a civil society and its economic competitiveness, as well as the success of individual Canadians, will hinge on having the best possible education and access to lifelong learning opportunities. Around the world, online learning—the use of digital networks to deliver and support learning opportunities—has emerged as a powerful and transformative means to meet these learning needs, as well as to extend and enrich traditional modes of instruction, at the post-secondary level. (p3)

## **E-Learning**

Like many other concepts and innovations in ICT, e-learning does not have a rigid definition and many authors and researchers have defined or described it in many ways. E-learning is networked online learning that takes place in a formal context and uses a range of multimedia technology (Garrison and Anderson, 2003). It has also been described as many different learning approaches that have in common the use of ICT (Clark, 2004). Murray has given it a simplistic description as being computer-facilitated instruction and learning oriented interaction (Murray, 2007). Summarily, e-learning may be described as a modern learning technology that is used to communicate or transfer knowledge, or perform teaching and learning activities interactively from afar with the aid of ICT tools. The use of technology in teaching and learning goes back centuries. What is however unique about e-learning is its multi-modal methodologies. It incorporates different media such as text, audio, and video. It can also be synchronous or asynchronous. E-learning has come to revolutionize the way traditional distance learning was conducted and it is indeed learning 'san frontier'.

Weller (2002) identified social acceptance, educator proximity, generic interface, interaction and personalization, and disruptive and sustaining technology as five reasons why e-learning has gained wide acceptance and likely to continue for a long time to come unlike other learning technologies that have fizzled away with time.

Besides reasons advanced by Weller, e-learning has other characteristics that makes it to stand out among all other learning technologies. E-learning is location independent as well as content independent. Knowledge can therefore be sourced from and taken to any corner of the world no matter how remote using appropriate network technology. It can also be time independent particularly in asynchronous form. Cost is an important factor in deploying any form of technology to the rural community. E-learning can attain a wide reach at low cost if implemented and managed properly. It can also be maintained at minimal cost.

Teaching processes can be automated, thus minimizing the effect of lack of qualified teaching personnel. Informal training can be maximized giving rise to new opportunities for non-professionals to find new roles.

There are several modes or forms of e-learning that can broadly be categorized into two; namely synchronous and asynchronous.

Synchronous e-learning is sometime referred to by other names like web conference, webinars, virtual classroom and online presentation among others. What all these names have in common however, is the use of computer network and software to facilitate live, interactive learning transaction via the internet. In synchronous e-learning, learning takes place in real time, that is, instantly and interactively between the teacher and students. It is usually scheduled and time-specific.

In asynchronous e-learning, collaboration, communication or access takes place intermittently. It is self-paced and available at any time. Murray (2007) outlines distinctive features and some examples of synchronous and asynchronous e-learning as shown in Table 3.

**Table 3** Synchronous vs Asynchronous e-learning (Murray, 2007: 2)

	<i><b>Distinctive features</b></i>	<i><b>Examples</b></i>
<b>Synchronous e-learning</b>	<ul style="list-style-type: none"> <li>▪ Real-time</li> <li>▪ Live</li> <li>▪ Usually scheduled and time-specific</li> <li>▪ Collective and collaborative</li> <li>▪ Simultaneous virtual presence of teacher and learners</li> <li>▪ Concurrent</li> </ul>	<ul style="list-style-type: none"> <li>▪ Instant messaging</li> <li>▪ Online chat</li> <li>▪ Live Webcasting</li> <li>▪ Audio conferencing</li> <li>▪ Video conferencing</li> <li>▪ Web conferencing</li> </ul>
<b>Asynchronous e-learning</b>	<ul style="list-style-type: none"> <li>▪ Intermittent access or interaction</li> <li>▪ Self-paced</li> <li>▪ Individual or intermittently collaborative</li> <li>▪ Independent learning</li> <li>▪ Usually available any time</li> <li>▪ Recorded or pre-produced</li> </ul>	<ul style="list-style-type: none"> <li>▪ Email</li> <li>▪ Threaded discussion</li> <li>▪ Boards</li> <li>▪ Web-based training</li> <li>▪ Podcasting</li> <li>▪ VCD or DVD</li> <li>▪ Computer-based training</li> </ul>

**E-Literacy – Bridging the digital divide**

(WSIS, Geneva 2003) envisioned the Information Society identified as one which is people centred, inclusive and development-oriented and where everyone can create, access, utilize and share information and knowledge. To achieve this vision of the Information Society, the WSIS Plan of Action recommended to:

- Establish rural information, communication and technology (ICT) access points
- Empower communities, especially those in rural and underserved areas, through the use of ICTs
- Promote distance learning, training and other forms of education as part of capacity building programmes
- Promote international and regional cooperation for capacity building

Even though, ICT penetration has increased significantly in developing countries in the last decade (see figure 1), a significant gap still exists between developed and developing countries. Also, the digital divide within countries and between different communities need be addressed if ICTs are to make meaningful impact and deliver on its promise.

It is, however, impractical to think of bridging digital divide in the rural communities without first finding ways to raise the low literacy level. Merely providing access to computers and the internet without the knowledge on how to utilize them will be a fruitless effort. It will also be difficult to persuade adults to leave their farms and other means of livelihood for the classroom in a conversational learning environment when some parents even see little reason to allow their children to go to school in the morning instead of helping out on the farm. As such, the only viable option is to design a kind of learning environment that will be flexible to the community life and schedules. Even in developed

economies such as the USA, balancing income earning with education has not been easy. Brogan, (1999) had observed that 50% of post-secondary students are working adults—over 6.1 million students.

“Edumocracy” or democratization of education—bringing instruction to people where and when they need it can only become a reality through the potentials and power of technology. Today’s information society and knowledge-based economy which are aspirations of all developing nations require a well-skilled human resource base. Today, we must understand systems in diverse contexts, and collaborate locally and around the globe using new tools like the Internet (Reich, 1995; Merrill Lynch, 1999). These attributes of learning contrast sharply with the low-level skills, content, learning and assessment methods that traditional ways of learning favor. The belief that learning takes place predominantly in the classroom and predominantly in children and youth is obsolete (Knowles, 1975, 1984; Cross, 1998).

Knowles’ (1984) work on “andragogy” – how adults learn ( in contrast to pedagogy—referring to children’s learning) evoked new discussions about how and why people learn, along with challenging assumptions that adults’ brains are not able to learn continually throughout life.

In practical terms, andragogy means that instruction for adults needs to focus more on the process and less on the content being taught. Knowles recommended strategies using tools like case studies, role-playing, simulations, and self-evaluation for adult education. Instructors would then play a role of facilitator or coach, rather than lecturer or grader.

These characteristics of adult education further strengthen the case for teaching and learning using the web in raising literacy level in rural communities. The opportunity to build new collaborative learning environments by using the web in teaching and learning is warranted based on the research that shows that students actually learn better while doing, and learn more when they are actively engaged in their learning (Brogan, 1999; Ussiph, 2008). ICT can play a key role in facilitating this collaboration.

## **Telecenters**

Telecenters or Community e-Centers (CeC) have gained prominence as a way of promoting and disseminating ICT knowledge in rural communities. Originating in Sweden around 1985, the concept of CeC is provision of a facility that provides public access to ICT-based services and applications for education, personal, social and economic development. The concept has now expanded to include telecentres, telecottages, community technology centres, and community communication shops, village knowledge centres, networked learning centres, multipurpose community telecentres (MCT), community access centres, and digital club houses.

CeC provide opportunities for development through ICT; extend the reach of public services such as education, health and social services; provide information of interest to the local community including farmers, local businesses and NGOs; and create new enterprises and job opportunities. The key function of a CeC is to provide public access.

CeCs are often equipped with telecommunications facilities (telephones, faxes), office equipment (computers, CD-ROMs, printers, photocopiers), multimedia hardware and software (radio, TV, video) and location for meetings and training

Services provided by CeCs include:

- Communication (telephone, fax, e-mail, Internet and radio)
- ICT training
- Agricultural information dissemination
- E-learning
- Distance health care
- E-government services
- Small and micro-enterprise support

There is no generally accepted model or optimal size for CeCs and every CeC operates with different objectives, services and ICT applications under different conditions based on the needs of the community. There are four different models of CeC – the adoption model, the government model, the commercial model and the school model.

Computers at CeC are networked with each one having an internet connectivity using a router via mostly dial-up connection and in some cases wireless modems. In the school model, a facilitator is assigned to the center that provides lessons on computer literacy and introduction to the Internet, usually in the afternoons and evenings for registered learners and also sees to the maintenance of the equipments.

With adequate resources and support, these telecenters can be transformed into a higher level, if not advanced open learning centers and institutions that will cater for the community's knowledge needs.

## **Research Methodology**

Action Research has been chosen as the most appropriate methodology for this research project. Dick, (2002) cited O'Brien, (2001) and McNiff (2002) that "Action research involves utilizing a systematic cyclical method of planning, taking action, observing, evaluating (including self-evaluation) and critical reflecting prior to planning the next cycle." (Dick, 2002: 2)

The actions have a set goal of addressing an identified problem in the workplace, for example, reducing the illiteracy of students through use of new strategies (Quigley, 2000). Action research may be viewed as a collaborative method to test new ideas and implement action for change. It involves direct participation in a dynamic research process, while monitoring and evaluating the effects of the researcher's actions with the aim of improving practice (Dick, 2002; Checkland & Holwell, 1998; Hult & Lennung, 1980). At its core, action research is a way to increase understanding of how change in one's actions or practices can mutually benefit a community of practitioners (McNiff, 2002; Reason & Bradburym, 2001; Carr & Kremmis 1986; Masters, 1995).

Action research may also be described as the pro-active application of tools and methods from social and behavioral sciences to practical problems, with the dual aims of improving practice as well as contributing to theory and knowledge in the area being studied. As words are the common currency in most activities of action research and since words produce qualitative information, action research therefore tends to be qualitative (Dick, 2002).

Often times action researchers either participate directly or intervene in a situation or phenomenon in order to apply a theory and to evaluate the value and usefulness of that theory. When undertaking action research the researcher starts with planning, continues to execution (intervention), observatory and reflection, before returning to planning and a new cycle begins. Remember, that we said it is a cyclical method of planning, taking action, observing and evaluating. Accordingly, the first cycle consists of the three phases:

### (a) Survey

A suitable research site will be chosen. It will consist of a school serving a local community in a remote location, lacking good access road and telephone access. The cooperating and active participation of school official and community leaders will be essential to success, as such careful selection of the site and sensitive briefing of residents will be required. Liaison with authorities will also be necessary. It is critical for the project to achieve an environment which is conducive to its experimental and participatory nature. Accordingly, once the site is selected, a study will be made of the school and of the social characteristics of the community a result of its having access to computers and the internet. Data gathered will relate to

- (i) demographic data and its dynamics
- (ii) educational levels
- (iii) commercial environment
- (iv) health matters

- (v) social needs and
- (vi) the cultural context.

(b) Implementation

The school will be equipped with appropriate technology for operation as a smart school. Training in its use will be provided to staff and pupils. Appropriate application will be introduced in line with published smart schools objectives. The equipment will be operated for sufficient time to allow a reasoned assessment of its effectiveness to be made.

The facility – telecenter – will be made available to the wider community out of school hours and used to pilot other internet applications for beneficial use of the computing and communication facilities now available to the community.

(c) Review

The results of the implementation activity will be reviewed in terms of whether or not there has been improvement to the delivery of teaching and learning, and academic performance of pupils during the period of implementation as well as direct benefits which the community enjoys as a result of gaining access to the internet. All effects will be highlighted and local views will be represented. The potential for wider impacts will be extrapolated for input to the continuing plans for further development. As part of the Action Research cycle, after suitable reflection, the research focus will return to planning and eventually into a new cycle of action.

### **Project Outputs**

- Specification of Information and Communication Technologies (ICTs) that will be capable of providing internet access to the remote communities of the designated school(s)
- Guidelines for the effective operation and maintenance of these technologies in order to achieve sustainable, reliable and acceptable performance.
- Processes for the application of these technologies for implementing the smart school concept in remote community schools
- Implementation plans and recommendations along with justifications and expected benefits, for the wider use of ICTs among rural community schools with the purpose of alleviating the worst effects of geographical isolation and economic marginalization suffered by rural communities.

### **Technology Transfer / Diffusion**

- Collaboration and partnership with education authorities, government agencies and NGOs can be utilized as a means of highlighting the need and opportunities derivable from Internet access by remote communities to exploit the opportunities and to service their needs.
- Market intelligence can provide a framework for the extension of commercially-viable internet services to remote villages in which Internet Service Providers can extend their services across regions of the country.
- Synergistic partnerships with technology suppliers can be used to open up investment opportunities in commercial applications which can serve as pump-primers for socially-oriented uses of the internet

## Direct Project Beneficiaries

- Pupils of schools in remote areas that constitute the majority of the Ghanaian rural communities. It is expected that the smart schools concept which is learner-centered educational services will become available to these pupils irrespective of where they live.
- Teachers in schools in remote rural communities
- Service providers who support operations of the remote rural schools
- Remote communities who will be able to exploit access to the internet for a variety of personal and commercial correspondences and also governmental information or matters.
- Politicians will have the opportunity of disseminating information of activities on local websites and receive comments and opinions of constituents via emails. This will be a boost for the e-government program.
- Suppliers seeking to engage in commercial or extension services with remote communities.

## Conclusion

This research is on-going and what has been presented in this paper is the review of the literature and issues regarding the use of ICT in education as well as efforts in other parts of the world to bridge digital divide in "endangered" rural communities. Outlined also is the research methodology being applied, technology diffusion, direct beneficiaries and expected output of the project.

It will worth mentioning here, that this research is a testing out kind – the replication of similar projects conducted elsewhere but in a different environmental setting to examine the output. The results or output of this research will be a major contribution to body of knowledge in the area of community informatics and the use of learning technology in supporting education and bridging the digital divide in rural communities.

## References

- Brogan, Pat (2006). Using the Web For Interactive Teaching and Learning – The Imperative for the New Millenium. Macromedia White Paper on E-Learning
- Bruce, W. D. et al, Technology Mediated Learning: Building Capacity In Rural Communities, The Journal of Community Informatics, Vol 3, No 4 (2007) available at <http://www.ci-journal.net/index.php/ciej/article/view/334/381>
- Dick, B. (2002) Action Research: Action and Research [On line]. Accessed Aug. 28, 2008 from <http://www.scu.edu.au/schools/gcm/ar/arp/aandr.html>
- Garrison & Anderson (2003). E-Learning in the 21<sup>st</sup> Century – A Framework for Research and Practice. RoutledgeFalmer Publishers, USA.
- Gurstein, M. (2007) What is Community Informatics and Why Does It Matter [On line] located at [http://eprints.rclis.org/archive/00012372/01/WHAT\\_IS\\_COMMUNITY\\_INFORMATICS\\_reading.pdf](http://eprints.rclis.org/archive/00012372/01/WHAT_IS_COMMUNITY_INFORMATICS_reading.pdf)
- Harris, W. Roger (2003) ICT For Poverty Alleviation, UNDP-APDIP, 2003, is located at <http://www.apdip.net/publications/iespprimers/eprimer-pov.pdf>
- Hyder Karen et al. (2007) The eLearning Guild's Handbook On Synchronous e-Learning [On line] located at [www.elearningguide.com](http://www.elearningguide.com)
- Kulik, C. L. and Kulik, J. A. (1986). Effectiveness of Computer-Based Education In Colleges. Association of Educational data Systems Journal, 19, p81 – 108

- Jhunjhunwala, A. n-Logue: The Story of a Rural Service Provider In India. *The Journal of Community Informatics*, Vol 1, No 1 (2004)
- Lie, E (2004) ITU/KADO Symposium On Building Digital Bridges: The Case of Malaysia is located at <http://www.itu.int/digitalbridges>
- Myers, M.D. (1997). *Qualitative Research in Information Systems*. *MIS Quarterly* (21:2), June 1997, pp. 241-242. *MISQ Discovery*, archival version, June 1997,
- Ramirez, R et al (2004), *Harnessing ICTs: A Canadian First Nations Experience* [On line] Accessed Oct. 20, 2008 from <http://www.smart.knet.ca>
- Reynolds, T, (2004) ITU/KADO Symposium On Building Digital Bridges available at <http://www.itu.int/digitalbridges>
- Sciadas, G (2004) *International Benchmarking For The Information Society*. ITU/KADO Symposium On Building Digital Bridges available at <http://www.itu.int/digitalbridges>
- Stringer, E. (2004). *Action Research in Education*. Upper Saddle River, NJ: Pearson Education, Inc.
- Tonio, V. L. *ICT In Education* [On line] Accessed Sept. 16, 2008 from [http://en.wikibooks.org/wiki/ICT\\_in\\_Education](http://en.wikibooks.org/wiki/ICT_in_Education)
- Walmart, B. *KiHS: Bridging the Traditional and Virtual Classroom in Canada's First Nation Schools*. *The Journal of Community Informatics*, Vol 1, No 3 (2005 p. 105)
- Weller, M (2002). *Delivering Learning on the Net – The What & How of Online Education*. Routledge Falmer Publishers, UK



# An Uncertain Future for Urban Vegetation

**Nicola Jane Wallbank and Philip James**, *School of Environment and Life Sciences, University of Salford*

## Introduction

The human population benefit greatly from nature (Bolund and Hunhammar, 1999) and incorporating more green areas into urban environments is a widely discussed topic. However, despite the advantages experienced from nature, human activities are continually posing a threat to natural ecosystems (Alberti, 2005). Vegetation is thought to be a key component to achieving sustainability (Van den Berg, 2007 and Chiesura, 2004).

Vegetation plays a significant role in urban environments as it provides a number of benefits to the human population which derive from ecosystem services (Bolund and Hunhammar, 1999). These services can have both direct and indirect impact on humans (Dwyer *et al.*, 1992) and they affect both the ecological and social features of urban areas.

The aim of this paper is to discuss the role of and introduce some of the ecological and social issues surrounding urban vegetation. It will also present and discuss the preliminary results of botanical surveys conducted within the study area of Runcorn, a town in the North West of England. The paper concludes with a discussion of the need for future work which combines these data with climate change predictions and with perceptions of local residents to develop scenarios for the development of this vegetative framework to *circa* 2060.

## The role of vegetation

Understanding the role of vegetation within built environments will help assess the external forces that influence its function and examine how it affects quality of life for the human population.

Vegetation within urban areas can provide a number of ecological functions such as air filtration, water and climate regulation and habitat provision (Norberg, 1999). Plants can filter air by physically trapping particles such as dust from the atmosphere and absorbing harmful gases for example; nitrogen dioxide and sulphur dioxide (Beckett *et al.*, 1998). Vegetation can retain and store water; this allows for natural drainage which reduces surface run off and therefore prevents flooding (de Groot *et al.*, 2002). Vegetation can regulate climate (Bolund and Hunhammar, 1999). It can reduce effects of wind and provide shelter during both warm and cold weather (NUFU, 1998 and Dimoudi and Nikolopoulou, 2003). These services help to control the environmental quality of towns and cities which reduces the need for engineered structures (Pickett and Cadenasso, 2008). Vegetation in the urban environment provides habitats for other wildlife species and therefore helps support and enhance biological and genetic diversity (Attwell, 2000).

Urban vegetation also provides direct benefits to humans. Recreational and psychological benefits in particular have received increased recognition over the last decade (Croucher, 2007). This is a result of the association between quality of life and the natural environment. The aesthetic value of vegetation can enhance residents' satisfaction, attachment and sense of responsibility thereby improving their overall well being (Groenewegen *et al.*, 2006). Visual encounters of vegetation have a positive influence on the human emotional and physiological state (Swanwick *et al.*, 2003). According to Parsons *et al.* (1998), people who drive through environments dominated by buildings have higher stress levels than those who drive mainly through natural environments. People who visit areas of green space regularly are less likely to experience stress-related illness than those who visit them less often (Grahn and Stigsdotter, 2003). Urban greenery can also provide opportunities for education and encourage physical activity (De Vries *et al.*, 2003). Vegetation can reduce noise pollution; only very dense vegetation cover can physically diminish noise levels, but there is evidence to suggest that screening the source of noise by trees can reduce the perceived noise level (Anderson, 1984 and Bolund and Hunhammar, 1999).

Economic advantages can be gained from the vegetation. Many studies have found a positive correlation between property prices and a close proximity to green areas (Luttik, 2000 and Tyrväinen and Miettinen, 2000). According to the Commission for Architecture and the Built Environment (2004) the provision of quality urban green space can enhance a cities reputation for high quality living environments and urban governance. Urban green spaces provide employment opportunities, and they help attract businesses and tourists to an area (Swanwick *et al.*, 2003).

### **The urban environment**

Succession of plant species within urban environments is often very different to that experienced in non-urban areas due to the variable anthropogenic influences (Sukopp, 2004), many habitats are kept at early succession stages due to regular disturbances (Niemelä, 1999). Urban areas are also known to support higher plant diversity than neighboring rural areas (Millard, 2002). This is the result of a high number of non-native species and the heterogeneity typical of urban landscapes that creates diverse habitats (Niemelä, 1999).

Fragmentation, isolation and adjustment to energy flow and nutrient cycling are common within urban areas (Alberti *et al.*, 2003). There are often difficulties such as soil compaction and herbicide applications (Sukopp, 2004). Humans significantly alter land cover and redistributed species which affects diversity, surface run off and pollution (Alberti *et al.*, 2003). Urban soils are also known to contain a high level of building materials such as cement which can increase its alkalinity (Sukopp, 2004). However, vegetation can also directly benefit from the physical attributes of the urban environment such as air movement and humidity levels (Attwell, 2000).

So urban areas can often be unique and difficult to predict. Therefore, to understand these ecosystems, the relationship between social and biological factors that control ecological processes must be studied.

### **Social intervention**

The inclusion and use of vegetation within urban areas is controlled by human perception. Social attitudes towards such vegetation are a complex issue, and one that is frequently contested. A variable that is likely to determine the perceived appropriateness of green space is the vegetation density (Bjerke *et al.* 2006). Vegetation is thought to facilitate crime, as it reduces visibility and can therefore conceal criminal activity. As a result densely vegetated areas often evoke feelings of insecurity (Kuo and Sullivan, 2001) and are avoided by those who perceive themselves as vulnerable to crime such as women, the elderly and children (Luymes and Tamminga, 1995). People often prefer landscapes that contain distant views and scattered tree grouping without the presence of understory layers rather than spaces containing dense trees and shrubs (Parsons, 1995).

Aesthetic and restorative functions attributed to urban green spaces can often be lost in environments that create a high fear of crime (Luymes and Tamminga 1995). As a result it is evident that the way in which vegetation is perceived can significantly affect part of its function.

The most influential criticism of naturalistic landscapes is their appearance (Özgüner *et al.*, 2007). However, people do appreciate the importance of nature and admire its beauty and Jorgensen *et al.* (2007) found that people wanted the option of visiting wild landscapes but did not want to live within them. They therefore concluded that people would prefer to have well-maintained landscapes in the immediate vicinity of their homes but want access to wilder areas within the local area.

Those who manage urban green spaces have the obligation to meet high demands for land-use and to create areas that meet the social, economic and environmental needs of the local community. One of the difficulties associated with the management of urban green space is the issue of finance. Ecosystem services are often undervalued (de Groot, 2006) and it is often the case that short term management options at low cost are preferred to costly long term strategies (Germann-Chiari and Seeland, 2004). A study by Özgüner *et al.* (2007) found that professionals recognised the importance of naturalistic environments but had voiced a concern over the required maintenance and heavy cost during the development stages.

## Botanical Survey of Runcorn

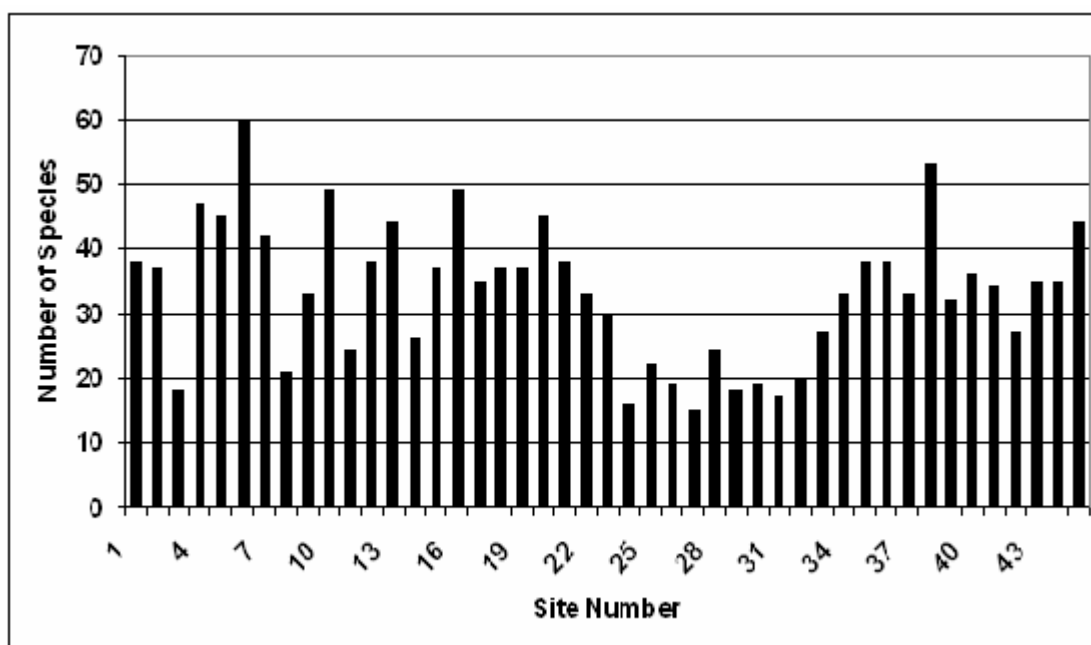
### Methods

The study was carried out in Runcorn, a town in North West England, situated 22km south east of Liverpool, on the south bank of the river Mersey. During the development of Runcorn existing topography and vegetation was complemented by large scale landscaping and tree planting. As a result Runcorn has a comprehensive open space system which includes playing fields, parks, woodlands and greenways; these areas form a vegetative framework surrounding all residential and commercial development. This framework comprises areas known as "structure planting". These vegetative areas act as wind breaks and noise barriers, improve aesthetic conditions within the town and provide opportunities for recreational activity. The trees and shrubs planted have now reached maturity and questions are being asked about their future. This provides the opportunity for gaining insights into the ecological and social issues associated with urban vegetation.

Floristic data will be collected from March 2009 to March 2010. All ground flora, shrub and tree species are considered. Forty five sites from across the whole town were randomly selected and surveyed. A detailed species inventory is being compiled. Moisture, pH and compaction readings have also been taken from each site.

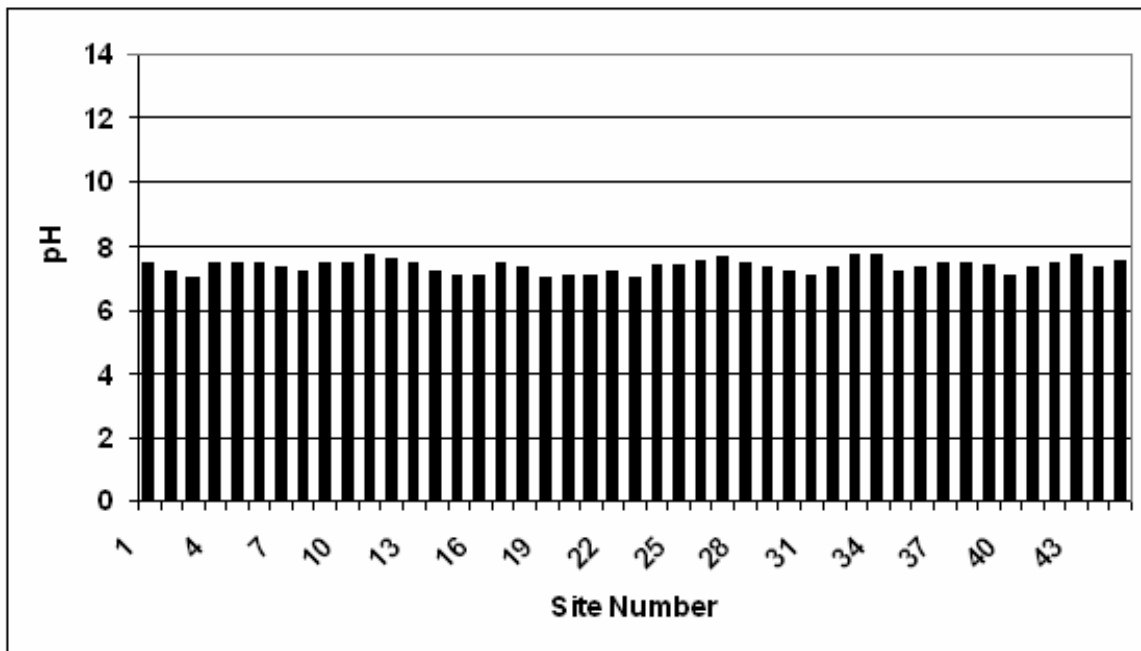
### Results and Discussion

Data presented here was collected at each of the forty five sampling sites between March and August 2009. Further data will be added to complete a twelve month sampling cycle. However, the months represented here are the main growing months in the UK and it is anticipated that few new species will be identified in the remaining months of the sampling period. The following figures allow comparisons to be made between each individual site regarding their species composition and environmental variables. There is a relatively high variation amongst each site and the species richness values range from 16 to 60 (Figure 1). A total of 138 species from 44 families have so far been recorded (24 tree, 15 shrub and 99 ground flora species).



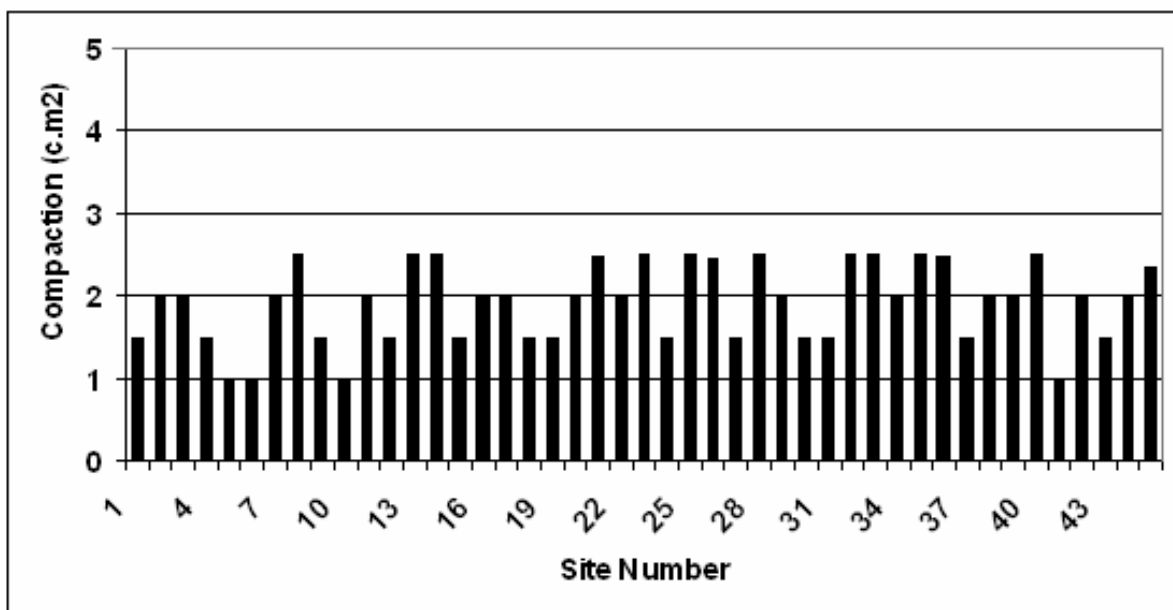
**Figure 1** Species richness values for each site.

All 45 sites had pH levels within a range of 7 to 7.9 which mean the soils are neutral to alkaline (Figure 2). The pH levels present at these sites fit within the typical range for UK soils and the limited range suggests that the whole of Runcorn is relatively homogenous with respect to this soil characteristic.



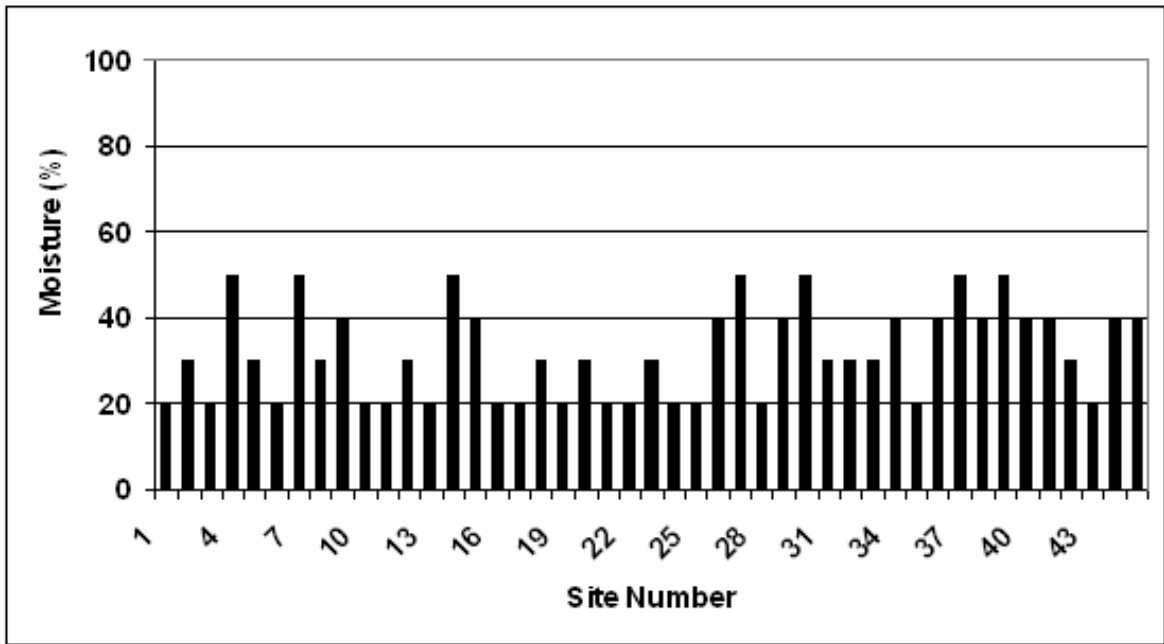
**Figure 2** pH values for each site.

Compaction levels for each site ranged from 1 to 2.5 (Figure 3). Areas of grassland typically have a high compaction reading (greater than 3) but the majority of 'structure planting' sites contain high amounts of shrub and tree layers which restrict the presence of grass species this may, therefore, explain the relatively low compaction readings.



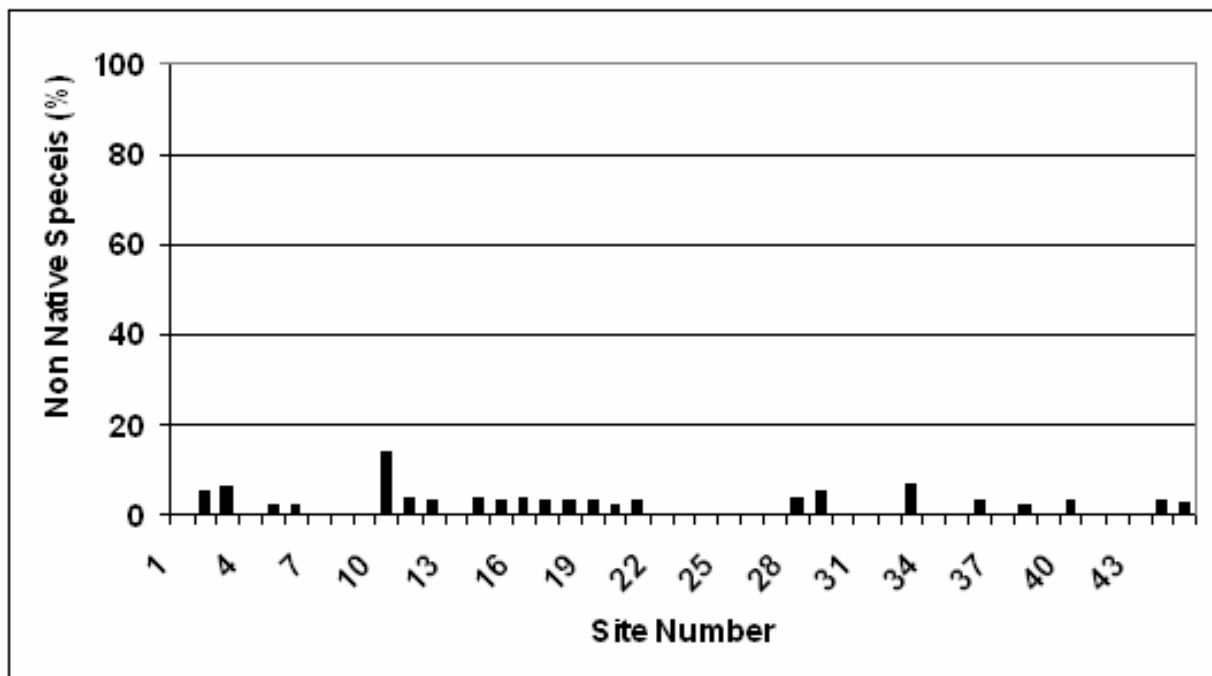
**Figure 3** Compaction levels for each site

Moisture readings ranged from 20 to 50% (Figure 4). A number of the 'structure planting' sites contained a wood chipping surface which could explain why a large majority of the sites showed a high level of moisture as this mulch would contribute to the retention of moisture within soils.



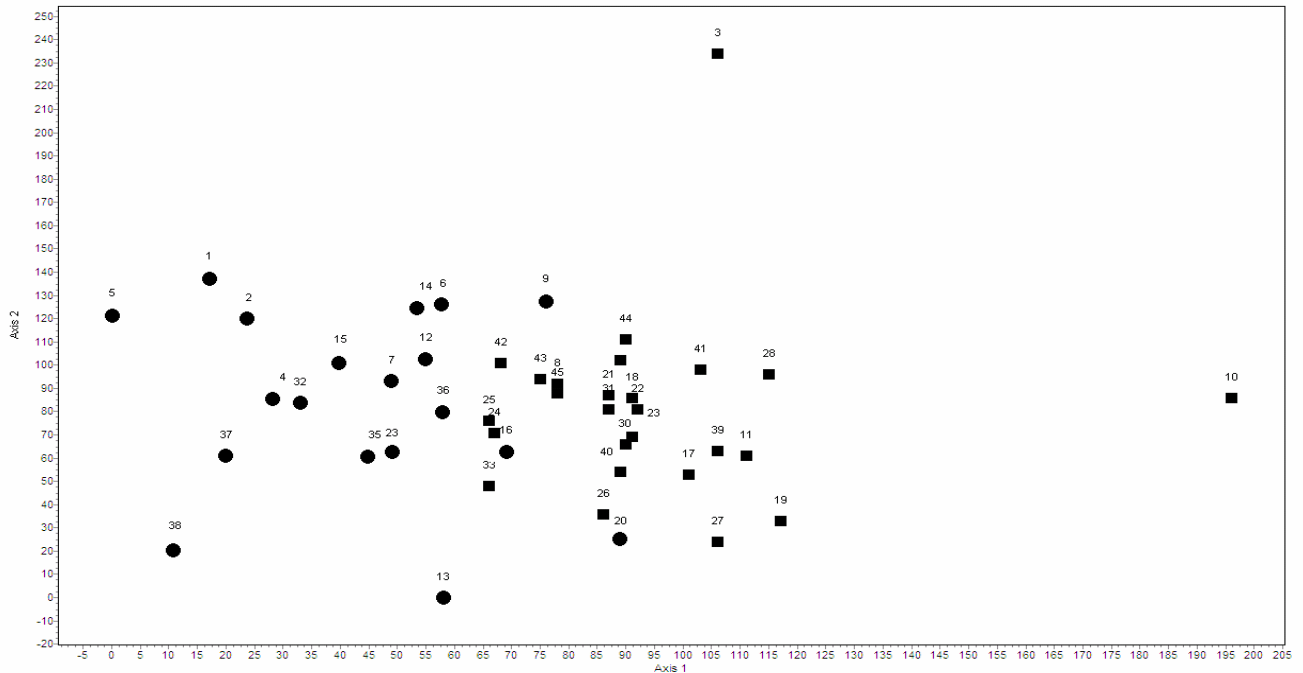
**Figure 4** Moisture reading for each site.

The percentage of non-native species present at each site ranged from 0 to 14% (Figure 5). Site 10 contained the most non native species with a total of 14% but 22 sites contained no non native species at all. The low number of non natives, whilst indicating that some plants have spread into these areas, suggests that the current species composition may not differ significantly from that which was originally planted.



**Figure 5** Number of non-native plant species at each site.

A DECORANA ordination plot was used to illustrate the similarity of the species composition between each site. The closer two sites are to each other in Figure 6 then the greater the similarity between these two sites. The numbers of each point represent the numbers assigned to each sampling site. Circular points represent sites that have a recreational function (group 1). Square points represent sites that have an aesthetic function (group 2).



**Figure 6** DECORANA ordination plot. Circular points represent sites that have a recreational function (group 1). Square points represent sites that have an aesthetic function (group 2). Numbers correspond to the individual sites.

Figure 6 demonstrates a close similarity between the majorities of the 'structure planting' sites, with the exception of sites 3 and 10. On closer examination of these two sites, it appears that site 10 contained the highest number of non-native species, as demonstrated in Figure 5. Site 3 contained only one tree species which is a large contrast to the other 44 sites. Very little variation is shown between the sites in relation to their pH, moisture and compaction levels, this could explain why the species compositions of the sites have remained similar since they were planted.

Further analysis was carried out to establish if the sites could be grouped based on other characteristics. Observations were made during the data collection of a high variation between each site and their function to the local population. Therefore all sites were divided into two different groups. One group contained the sites that have a direct interaction with humans, such as greenways and parks and the second contained sites that receive little human disturbances such as those used for aesthetic purposes. Figure 6 demonstrates that sites within group 1 have a closer similarity in species composition to each other than those within group 2, and *vice versa*.

Varying levels of management intervention has also been observed. For example, the two sites that contain the highest number of species, sites 5 and 38, are both large in size but differ significantly in management. One receives very little intervention and has developed into a wildflower meadow the other receives regular maintenance practice and has the appearance of a park.

The similar species compositions within all areas of structure planting suggests that in the future only one set of management options could be produced and applied to all sites. However, further research is required to establish the abundance of different plant species.

To sustain the ecosystem services attributed to urban vegetation assessments must be made on the external factors that are likely to cause change. A growing concern for the future of urban green space is the effect of climate change. Anticipated changes to vegetation include; alterations to phenology, the arrival of non-native species and changes to community compositions (Walther *et al.*, 2002). Furthermore climate change will encourage the relocation of certain species due to changes in the geographical distribution of suitable habitats (Hill *et al.*, 1993). Future work will include the combining of the ecological data with climate change projections.

Vegetation plays many roles but there are some less desirable associations. The literature regarding peoples' perception suggest that people want contact with nature but that some forms of vegetation such as very dense planting can provoke feeling of insecurity. Therefore further work is required to understand in greater detail just what people want from urban green spaces.

It is also evident that management of urban vegetation can differ significantly therefore the different management trends and the reasons for applying varying intervention plans require further understanding.

The authors of this paper have presented preliminary findings from a three year study. The data presented has been collected within the first 11 months of this study. Over the next two years these data will be explored more thoroughly and combined with data recording the current management practices amongst landowners within Runcorn, and the impacts of predicted climate change in order to inform land owners with respect to future management strategies for these important areas of natural environment within the town.

## References

- Alberti, M. (2005) 'The effects of urban patterns on ecosystem function', *International Regional Science Review*, **28**, 168-192.
- Alberti, M., J. M. Marzluff, E. Shulenberger, G. Bradley, C. Ryan & C. Zumbunnen (2003) 'Intergrating Humans into ecology: opportunities and challenges for studying urban ecosystems', *BioScience*, **53**, 1169-1179.
- Anderson, L. M., Mulligan, B E, Goodman, L S (1984) 'Effects of Vegetation on Human Response to Sound', *Journal of Arboriculture*, **10**, 45-49.
- Attwell, K. (2000) 'Urban land resources and urban planting -- case studies from Denmark', *Landscape and Urban Planning*, **52**, 145-163.
- Beckett, K. P., P. H. Freer-Smith & G. Taylor (1998) 'Urban woodlands: their role in reducing the effects of particulate pollution', *Environmental Pollution*, **99**, 347-360.
- Bjerke, T., T. Østdahl, C. Thrane & E. Strumse (2006) 'Vegetation density of urban parks and perceived appropriateness for recreation', *Urban Forestry & Urban Greening*, **5**, 35-44.
- Bolund, P. & S. Hunhammar (1999), 'Ecosystem services in urban areas', *Ecological Economics*, **29**, 293-301.
- Chiesura, A. (2004) 'The role of urban parks for the sustainable city', *Landscape and Urban Planning*, **68**, 129-138.
- Communities and Local Government (2008) *Draft Planning Policy Statment: Eco-towns-consultation* [online] available at:  
[www.communities.gov.uk/publications/planningandbuilding/ppsecotowns](http://www.communities.gov.uk/publications/planningandbuilding/ppsecotowns) (accessed 27 Augst 2009).
- Croucher, K., Myers, L., Bretherton, J. (2007) The links between greenspace and health; a critical literature review. Unpublished.
- de Groot, R. (2006) 'Function-analysis and valuation as a tool to assess land use conflicts in planning for sustainable, multi-functional landscapes', *Landscape and Urban Planning*, **75**, 175-186.
- de Groot, R. S., M. A. Wilson & R. M. J. Boumans (2002) 'A typology for the classification, description and valuation of ecosystem functions, goods and services', *Ecological Economics*, **41**, 393-408.
- De Vries, S., R. Verheij, A. , P. P. Groenewegen & P. Spreeuwenberg (2003) 'Natural environments-healthy environments? An exploratory analysis of the relationship between green space and health', *Environment and planning* **35**, 1717-1731.
- Dimoudi, A. & M. Nikolopoulou (2003), 'Vegetation in the urban environment: microclimatic analysis and benefits', *Energy and Buildings*, **35**, 69-76.
- Dwyer, J. F., E. G. McPherson, H. W. Schroeder & R. A. Rowntree (1992), 'Assessing the Benefits and Costs of the Urban Forest', *Journal of Arboriculture*, **18**, 227-234.

- Germann-Chiari, C. & K. Seeland (2004), 'Are urban green spaces optimally distributed to act as places for social integration? Results of a geographical information system (GIS) approach for urban forestry research', *Forest Policy and Economics*, **6**, 3-13.
- Grahn, P. & U. A. Stigsdotter (2003) 'Landscape planning and stress', *Urban Forestry & Urban Greening*, **2**, 001-018.
- Groenewegen, P., A. van den Berg, S. de Vries & R. Verheij (2006) 'Vitamin G: effects of green space on health, well-being, and social safety', *BMC Public Health*, **6**, 149.
- Hill, H. O., P. D. Carey, B. C. Eversham, H. R. Arnold, C. D. Preston, M. G. Telfer, N. J. Brown, N. Veitch, R. C. Welch, G. W. Elmes & A. Buse (1993) 'The role of corridors, stepping stones and island for species conservation in a change climate', *English Nature Research Reports*. Peterborough: English Nature.
- Jorgensen, A., J. Hitchmough & N. Dunnett (2007) 'Woodland as a setting for housing-appreciation and fear and the contribution to residential satisfaction and place identity in Warrington New Town, UK', *Landscape and Urban Planning*, **79**, 273-287.
- Kuo, F. E. & W. C. Sullivan (2001) 'Environment and crime in the inner city, does vegetation reduce crime?' *Environment and Behaviour*, **33**, 343-367.
- Luttik, J. (2000) 'The value of trees, water and open space as reflected by house prices in the Netherlands', *Landscape and Urban Planning*, **48**, 161-167.
- Luymes, D. T. & K. Tamminga (1995) 'Integrating public safety and use into planning urban greenways', *Landscape and Urban Planning*, **33**, 391-400.
- Millard, A. (2008) Semi-natural vegetation and its relationship to designated urban greenspace at the landscape scale in Leeds, UK. *Landscape Ecology*, **23**, 1231-1241.
- Niemelä, J. (1999) 'Ecology and urban planning', *Biodiversity and conservation*, **8**, 119-131.
- NUFU (1998) *Trees matter- The benefits of trees and woods in Towns*, National Urban Forestry Unit.
- Parsons, R. (1995) 'Conflict between ecological sustainability and environmental aesthetics: Conundrum, canard or curiosity', *Landscape and Urban Planning*, **32**, 227-244.
- Parsons, R., L. G. Tassinary, R. S. Ulrich, M. R. Hebl & M. Grossman-Alexander (1998) 'The view from the road: implications for stress recovery and immunization', *Journal of Environmental Psychology*, **18**, 113-140.
- Pickett, S. T. A. & M. L. Cadenasso (2008) 'Linking ecological and built components of urban mosaics: an open cycle of ecological design', *Journal of Ecology*, **96**, 8-12.
- Pickett, S. T. A. C., M. L. Grove, J. M. Nilon, C. H. Pouyat, R. V. Zipperer, W. V. Constanza, R. (2001) 'Urban ecological systems: Linking terrestrial ecology, physical and socioeconomic components of metropolitan areas', *Annu. Rev. Ecol. Syst.*, **32**, 127-157.
- Sukopp, H. (2004) 'Human-caused impact on preserved vegetation', *Landscape and Urban Planning*, **68**, 347-355.
- Swanwick, C., Dunnett, N., Woolley, H. (2003) Nature, role and value of green space in towns and cities and overview. *Built Environment*, **29** (2), 94-106.
- Tyrväinen, L. & A. Miettinen (2000) 'Property Prices and Urban Forest Amenities', *Journal of Environmental Economics and Management*, **39**, 205-223.
- Van den Berg, A. E., Hartig, Terry., Staats, Henk (2007) 'Preference for Nature in Urbanized Societies: Stress, Restoration, and the Pursuit of Sustainability', *Journal of Social Issues*, **63**, 79-96.
- Walther, G.-R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J.-M. Fromentin, O. Hoegh-Guldberg & F. Bairlein (2002) 'Ecological responses to recent climate change', *Nature*, **416**, 389-395.
- Özgüner, H., A. D. Kendle & R. J. Bisgrove (2007) 'Attitudes of landscape professionals towards naturalistic versus formal urban landscapes in the UK', *Landscape and Urban Planning*, **81**, 34-45.



# **A Voice for the People: the role of Social Media technologies in Community Empowerment and Regeneration**

**Eileen Wattam**, School of Media, Music & Performance, University of Salford.

## **Abstract**

This paper examines whether and how social media technologies are linked with a greater potential for community empowerment and regeneration. The paper is based on the early stages of a qualitative study of a Community Reporters Programme being implemented by a not for profit community development organisation which seeks to empower communities by providing participants with a voice through social media technologies to tell stories about their lives and their local area. Given that many of these technologies are open source in nature the potential to overcome financial and social sustainability challenges faced by community-based initiatives is also briefly considered.

## **Introduction and Background**

The following provides a brief background and introduction to the research on which this paper is based by discussing the key concepts and relationships pivotal to the research and the identified need for further work which has driven the study. During the 1990s the relationship between Information and Communication Technologies became the focus of academic attention and was surrounded by much hope and rhetoric (Aurigi, 2005). At the same time a growing number of ICT focused community and social initiatives began to emerge across the world (Hunt and Stout, 1997; Graham, 1997; Graham and Marvin, 1999). The growing interest in the link between ICT, community development and regeneration are perhaps best illustrated in the emergence of Community Informatics as a field of study and practice toward the close of the 20<sup>th</sup> Century (Gurstein, 2000). In the UK, ICT and the concept of 'community' have become increasingly central to the urban policy agenda, ICT being seen as a vital tool in regenerating and empowering deprived communities (Sherman, 1999; Hellowell, 2001; PAT 15, 2000). From the late 1990s a range of social and community ICT initiatives began to emerge, many seeking to directly impact on regeneration. Surprisingly, however, the link between ICT and regenerating deprived urban areas remains a relatively new concept in the UK (Southern and Townsend, 2005) and there is a paucity of literature and research within this area. The early optimistic and deterministic portrayal of ICTs as somehow being inherently transforming and empowering have more recently been challenged by commentators who are beginning to consider whether such technologies are empowering and whether they can really be used to address the many complex issues involved in regenerating deprived areas. The ambiguity and value laden assumptions which surround the area are thus being brought to light suggesting that the area requires further investigation. The benefits associated with ICT remain unclear and as yet remain unsubstantiated and it is suggested, unmatched by the experiences of communities (Southern, 2002). In addition community-based ICT initiatives and ICT regeneration initiatives generally are described within the existing literature as highly problematic, not least because they operate within complex environments and suffer from funding and sustainability issues (Dabinett, 2001; Southern, 2002; Southern and Townsend, 2005). There are also very few empirical studies of ICT as a means for urban regeneration (Southern and Townsend, 2005).

## **ICT, Urban regeneration and the digital divide**

ICTs have traditionally been embraced by politicians and academics as a means to overcome existing social inequalities and much of the early discourse emphasises their capacity to empower people and communities (INSINC, 1997, PAT, 15, 2000; Hellowell, 2001; Selwyn, 2004). In the UK ICT and urban regeneration have been brought together in the context of addressing the digital divide driven by concerns which emerged among such optimism that existing social inequalities would be exacerbated if people and communities were excluded from the benefits of the information society. The idea that the role of ICT in urban

regeneration is now central to the national political agenda, it has been suggested, is nowhere more apparent than in the work of the Social Exclusion Unit's Policy Action Team 15 (Leach, 2000), a team which was commissioned to develop a strategy to increase the availability and take up of ICTs in deprived neighbourhoods. This led to a series of initiatives to promote access to ICTs with a particular emphasis on deprived communities. However, the evidence suggests that digital exclusion remains a real phenomenon (Sassi, 2005; Klecun, 2008) and an important area of social debate (Selwyn, 2004). It is also now increasingly recognised that the lens of digital divide needs to be expanded beyond notions of access to the technology, that new frameworks for understanding the complexities of the Digital Divide need to be developed and that a wider range of social resources beyond access are needed to promote social inclusion and empowerment through ICT (Sherman, 1999; Clement and Shade, 2000; Harris, 2000; Gurstein, 2003; Warschaucer, 2002; 2004; Selwyn, 2004; Sassi, 2005; Potter, 2006; Eubanks, 2008; Klecun, 2008; Zheng and Walsham, 2008). The latest research suggests that the link between digital exclusion and social disadvantage is persistent in the UK (Helspher, 2008), while others question the relationship between the two areas (Bure, 2005). Klecun (2008) suggests that the Digital Divide discourse is becoming more sophisticated over time and that it is now widely acknowledged within academia and policy that it is not as simple as a divide between the haves and the have nots, that it is more complex and in particular linked to social exclusion (Selwyn, 2004; Cushman and McClean, 2006, Potter, 2006; Klecun, 2008). As a result terms such as digital exclusion or digital inequality are also being used and it is argued are more appropriate than the term digital divide (Klecun, 2008). It is also increasingly recognised that inherent assumptions within the lens of the digital divide about people and technology need to be challenged (Gurstein, 2003; Selwyn, 2004, Potter, 2006; Wyatt, 2005). The digital divide, or perhaps more appropriately termed digital inclusion, as Klecun (2008) suggests, is thus a central concept to ICT and urban regeneration. Thus the need for further research in relation to ICT and the regeneration of deprived communities has been made clear as has the associated need to engage with discourses around digital inclusion, empowerment and social exclusion.

### **Social Media, Digital Inclusion and Empowerment**

While ICT is linked to urban regeneration, community development and empowerment, defining what is meant by ICT itself is a difficult task simply because there are so many technologies which could come under the umbrella term. As a result, Selwyn (2004) suggests that definitions provided by academics and politicians can be quite limited, vague and often either too broad, "*treating ICT as a homogenous concept*" (Selwyn, 2004: 346) or too narrow in focus. Social Media has been defined as broadly including social networking sites, blogs, user generated videos and pictures, message boards and wikis (Li and Bernoff, 2008; Shrock, 2009). By concentrating on social media, this research has a very specific focus in terms of an interest in the content dimension which ICT enables, often referred to as user generated content (Mork Peterson, 2008). The focus of the research also reflects the increasing convergence of new and old media enabled by digital technologies around web 2.0 and their tendency to permeate everyday lives, (Coopman, 2009). The research also seeks to move beyond traditional digital divide discourses which tend to focus on access to computers and the Internet to incorporate the consideration of the range of devices involved in content generation and production. Ultimately the focus on social media is driven by the commonly expressed association between such technologies and empowerment based on the premise that they enable people to be producers rather than passive consumers of content (Hellawell, 2001). Understanding the relational aspects of such technologies however reveals that content may be associated with different forms of value and significance, both positive and negative (Mork Peterson, 2008), suggesting that the area requires further exploration. Additionally recent research suggests that as the use of technology becomes more advanced – technologies such as social networking – so the barriers to digital engagement increase (Helspher, 2008).

### **Community empowerment and Social Media**

Community-led web-sites and networks have for some time now been strongly linked to empowerment (Miller, 1999) and it is the interactive uses of technologies which are seen as central to the potential for community and citizen empowerment (Loader and Keeble, 2004). Recently social and community media has become a strand of the policy response to digital

exclusion embodied in the digital mentors programme designed to help people and communities in deprived areas use social and community media (Department for Communities and Local Government, 2008). The research on which this paper is based can therefore make a valuable contribution to discourses which are bound to escalate around this dimension of digital inclusion. Additionally because the research focuses on a Community Reporters Programme the research is also relevant to discourses around technology and community media and journalism (Gillmoor, 2004). In addition, to date there is little research evidence to support a relationship between community empowerment and social media. Dara Byrne's recent research (2007), for example, which studies community life on a black social network, found that the potential for social mobilisation through community networks online is not yet realised.

## **Research Design and Methodology**

The research is qualitative in nature and predominately located within the interpretive research paradigm, viewing reality therefore as socially constructed and seeking to understand phenomenon through the meanings that people assign to them. The critical research paradigm is also drawn upon to recognise that the capacity of people to bring about change may be limited by various forms of domination (Myers, 2003). The methodology can be described as an interpretive case study, an approach seen as valuable in understanding the complexities of the links between social media, empowerment and regeneration through in-depth contextual research. The methods employed include interviews and participant and non participant observation. Evidence from various documents, websites and online social and community networks also inform the findings. This paper is thus based on a case study of a Community Reporter's Programme focused on two key regeneration areas in Manchester and Salford and developed by a not for profit organisation. The case study was chosen on the basis of having clear objectives with regard to digital inclusion, empowerment and regeneration. The findings presented within this paper are based on the very early stages of the research and therefore predominately draws on semi-structured interviews with 13 participants in the programme i.e. Community Reporters and interviews with six staff running and implementing the programme. The research is broadly located within the field of Science and Technology Studies thus rejecting technologically deterministic views of the world and recognising that technologies are socially constructed by the broader social, political, economic and cultural context (Bijker et al, 1987). This overarching view also seeks to avoid social determinism and recognises change as a socio-technical process, the technology itself therefore also being open to analysis (Van Lies Hout et al; 2001). This school of thought has thus informed the analysis and interpretation of the findings. This paper draws on a few examples of community reporters and projects within the overall case study to illustrate the complexities of the relationship between the areas of study.

## **Social Media and Contributions to community empowerment and regeneration: Findings**

### *The Community Reporter's Programme*

The Community Reporters Programme has been developed by People's Voice Media, a community development organisation in Manchester which uses social media to empower people and communities. The Community Reporters Programme seeks to give people a voice and build their confidence in the use of social media tools to enable them to express themselves and tell their own personal stories or stories about their local area online. Community Reporters use a range of social media tools including blogs, videos and podcasts and use various devices including video cameras, mobiles, voice recorders and cameras to capture and distribute a wide variety of content. While the programme is expanding this research is based on the initial focus of the programme which was on two Government driven regeneration initiatives i.e. New Deal for Communities regeneration initiatives in Salford and Manchester. Community reporters learn how to produce content with new social media tools such as blogs, podcasts and films and a range of devices including digital cameras, mobile phones and video cameras which they upload to the organisation's networks, on other community networks or through social networking sites such as you-tube and facebook. Community reporters interviewed produce a range of content including; video stories about experiences of being homeless; theatre and music reviews; short films about community events and organisations; podcasts about the work of a local artists, develop blogs, for

example to promote local music or history. People's Voice Media works with individuals through community reporter training sessions and also works with groups to help them integrate social media as a tool for reaching their objectives or simply to enable them to 'tell stories.' Community reporters are people who essentially attend training in community reporting and move on to volunteer as a community report for People's Voice Media. They may also be individuals who are already active within their communities. In order to understand how social media might have a role for community empowerment within this social-technical process it is firstly necessary to establish what is meant by the concept of community empowerment.

Community empowerment is a difficult concept because it is open to different interpretations by different people, generally is related to its understanding as either a product or a process (Laverack, 2001) and additionally community empowerment is a multidimensional concept since it has psychological, social, economic and political dimensions (Czuba and Page, 1999). However most often within the literature community empowerment is presented as a process which enables people to organise and mobilise so they can act on issues they find important to them (Laverack, 2001; Czuba and Page, 1999). A key feature of the concept is the link between individual and collective change, individual empowerment being viewed as a prerequisite for community empowerment in terms of people developing a spirit of community (Wilson, 1996). Individual empowerment is often linked to psychological empowerment commonly demonstrated in self confident behavior and is also presented as people being able to gain control of their lives (Czuba and Page, 1999, Laverack and Wallerstein, 2001; Friedmann, 1999). The evidence suggests that there is an association between social media and individual empowerment in terms of the development of psychological power and social power such as increased confidence, increased ability to interact with people, new contacts and friends. Other respondents reported feeling empowered by putting content online.

To illustrate the complexities of how social media features in individual empowerment the example of a community reporter called Kevin is utilised. Kevin joined the programme because he was looking for voluntary work that matched his interests and background in audio visual work. When he joined the programme he had been out of work for medical reasons for three years and felt that embarking on the programme would be a good way of getting back into work. Since joining the programme Kevin learnt how to create his own blog, a music blog with the aim of promoting Salford bands who would not typically get coverage in local newspapers. This led him to film a music night at a local pub and make a video of it. He has also learnt computer skills including being trained in how to use a video editing package and he has gone on to make several short video films and one audio piece which he has posted on one of the organisation's distribution channels. Kevin has attended several of the organisation's training sessions and has found the whole process really helpful socially in terms of boosting his confidence and reconnecting him with his local area and Salford. For example he said that; *"When I first moved into where I lived it was difficult to get to meet and talk to people, now when I go into the pub people are talking to me about the blog and it seems to be an easy way to talk to people because everybody likes music."* He also felt at the time of interview that the process has really improved his job prospects and since then he has gone on to gain some ad-hoc work with People's Voice Media. Interestingly Kevin also talked about going out to different locations and meeting different people as a key part of the benefits he has received which he also commented *"has changed my perception of the town and made it more interesting."* The example of Kevin will be related to in further detail in a later section on the socio-technical process which promotes or inhibits empowerment. While there are further examples of people like Kevin there are people who drop out of the process or of course never get involved, issues which are also discussed later in the paper.

The link between individual and community empowerment is perhaps already implied in the notion of community reporting yet some staff and community reporters feel that the process is definitely empowering for individuals but the link to a wider community benefit is less apparent. Analysis of the interviews with community reporters does suggest that to date the process does not seem to be having an effect in terms of people organising and mobilising for collective change, a key aspect of community empowerment. However for groups who are already organised there is some evidence to suggest that social media does provide a tool to assist them reach their goals. The most relevant example is a project called Edge Lane Allotments, which was saved from housing development by a group of activists who have been working to regenerate the site. People's Voice media have been working with the initiative to advertise the project. Patrick the plot steward has been trained as a community reporter and a

few short films have been made about the site. Patrick, for example, made a video appealing for people to take up the remaining plots, which resulted in groups and organisations coming forward to help the project in terms of volunteers to help to clear up the site which has fallen into disuse and suffered as a result of contractors beginning work to take over the site. Patrick, the plot steward said that *"it was an excellent way of broadcasting and contacting various groups"*. Such groups have pledged their support as a result in the form of, for example, assisting with clean up days. Various people on the allotment also now post pictures and information about the allotment on the organisations blog site they have developed within east Manchester. One member of the allotment groups said that through doing this it has really empowered her. A key aspect of the work with the groups has been one of the staff going out into the allotment site with the technology such as a laptop to provide guidance and support. Contributions of social media to community empowerment are also dependent on interpretations of 'community' which has long been argued is a difficult concept to define. People have individual ideas of the meaning of community and thus of the role of social media within that.

Edge Lane does however provide an example of how social media is linking to community empowerment in the context of regeneration. Community empowerment is often presented within policy as being about people being able to solve their own problems or people having a voice, particularly in terms of influencing decision making. The intersection between the two areas could therefore be described as lying within the political domain of community empowerment. Again a specific example is used here to illustrate how social media is featuring within this domain. Talk Broughton is a partnership regeneration project between housing providers funded by the City Council. The focus of the project is to assist Salford residents to access construction training and employment opportunities and as part of this research is being undertaken in the area of Higher Broughton in Salford with the purpose of looking at why many people from this area are inactive. The Community Reporters Programme has been integrated within this project training local people as community reporters to go out into the area and ask them about their views of the area, employment and services. The reporters thus went out armed with audio and video recorders interviewing local people and made a short video to summarise these views. One of the community reporters involved said that he felt that the video was a powerful way of presenting how residents really feel. The video was played at a launch event attended by decision makers. The actual impact on regeneration and change remains to be seen however. Certainly for some of the community reporters involved it has been a valuable process. Euna, for example, an asylum seeker from Zimbabwe was living in the area and really enjoyed going out interviewing local people. She wanted to develop her communication skills and found it valuable in this respect. Euna, for example, said *"I learnt a lot of skills, blogging, audio, communicating with people, I enjoyed myself a lot, going out interviewing people on the streets. I love it."* Euna said that not only has the process empowered her in terms of developing new skills but she also feels it has empowered the community by providing them with a voice, *"they know their voices can be heard, anyone can say anything regardless of their colour or whatever and they know their views are going to be heard, so it is good for the people of Broughton."*

However it is also interesting that some issues of concern do go unvoiced by community reporters. It may be expected that linking social media to community reporting with a focus on people living in regeneration areas would lead to people potentially voicing opinions on regeneration issues. Some community reporters spoken to were particularly concerned about housing issues locally, describing approaches as 'modern day slum clearance.' A further respondent said that; *"regeneration can be a euphemism for slum clearance."* Jane, a community reporter particularly in relation to community radio and recording the views of older people about the past said that *"from the people I speak to regeneration spells the end of community."* While one of the respondents said he would like to act as a voice for the people in highlighting issues like housing there was some identified reasons why people may not want to voice some issues despite strong feelings. One aspect is that community reporting does have quite positive interpretations in that people want to use social media to produce content about the good things going on in their area and not focus on the bad things linked to counteracting negative coverage of such areas within the mainstream press, others simply do not want to get too controversial or political. A couple of community reporters expressed their feelings that there is no campaigning element to the programme and that they feel that maybe there is an underlying fear of offending people and organisations particularly when the funding has come from the regeneration bodies locally. There are also other outlets which voice such

concerns such as the community magazine the *Salford Star*. The impact of the fact that the Community Reporters Programme is funded by local regeneration bodies on content that is being encouraged and produced is emerging as an area of interest which requires further investigation suggesting that the contribution of social media is influenced by issues relating the local political and funding context.

Overall social media, as evident through this discussion of the Community Reporters Programme, does provide people with an approaches to ICT and regeneration which have tended to focus on the economic imperative (Loader, 2002; Stoecker, 2005). It is also clear from the examples provided of various people and projects social media works or does work for community empowerment and regeneration as part of a socio-technical process of change. Drawing in the framework of domestication, the next section looks at how the social significance of social media in this context is shaped by various elements of social media appropriation followed by the key social actors.

### **Shaping the potential for community empowerment and regeneration,**

Drawing on the framework of domestication (Silverstone et al., 1992; Sorensen, 2005) and the examples discussed within this paper, the socio-technical elements which shape the potential for community empowerment and regeneration can be illustrated through considering practical, symbolic and cognitive dimensions of domestication. The practical elements shaping domestication include access to the social media centres, equipment and the distribution channels set up by the organisation. The role of community-based centres and community networks in fostering inclusion and regeneration has long been recognised (PAT 15, 2000) what does appear to be new is the borrowing of equipment which in practical terms enables community reporters who would not otherwise be able to, to go out into their communities with the equipment such as various recording devices. This actual process of community reporters going out in the local area with various devices is also interestingly a key element of empowerment. For example, Kevin found that the process helped him to get out of the house and reconnect with his local area. This in turn is indicative of the era of ubiquitous computing and the increasing mobility of technology. This does however raise new issues of safety and privacy. For example some community reporters felt they have to go out in teams to protect the equipment. The practical issues involved in filming and recording people also mean that for some it is better to go out as part of a team. This challenges views of social media appropriations as socially isolating, challenges dichotomous separations between the real and the virtual and points to the complexities of the appropriation of such technologies beyond use of the Internet and in particular use as a tool for social networking which much of the research has concentrated on to date (Boyd and Ellison, 2007).

Community reporters interviewed on the whole found the social media centres to be essential particularly in terms of making links and connections as well as borrowing equipment. Some practical issues which were voiced included equipment sometimes being of poor quality or being unavailable. The difference between the equipment available in East Manchester as opposed to East Salford is an issue of concern for people living in East Manchester who feel marginalised because the Salford Social Media have such a vast array of equipment in comparison. Practical issues of safety and security lie at the heart of the situation. The Grange in Manchester, for example, did incorporate top of the range equipment such as Apple Mac computers but they were all stolen within a few weeks. The Grange however has a much stronger community link physically being within a school and a housing area. Salford Social Media Centre in contrast is located within an innovation forum with strong security measures such as security gates and barbed wire. These practical measures to integrate technology could impact on accessibility to all, one respondent suggested.

Another practical issue concerns the distribution channels which community reporters are provided with to upload the content they produce. These come in the form of a network of some 14 community websites managed by the organisation. Two of these websites relate specifically to the regeneration areas in east Salford and East Manchester. Staff and volunteers moderate the sites and ensure that content does not break the guidelines set out in their editorial policy. Some community reporters expressed that they feel that sometimes content gets lost and goes to the wrong place, this leaves them frustrated when they have put a lot of effort into producing work, some people were even finding it difficult to find their own content. Also some people felt it was hard to relate to the boundaries of the regeneration areas on which the two community sites they are predominately encouraged to upload content on are

based. There was evidence for example of some confusion over what kinds of content relating to which areas are appropriate for these sites. Quantity of content was also a related concern expressed by a few people, with some concerns that the emphasis of the organisation is to put content online as quickly as possible, some feel this limits the in-depth examination of issues. Publicity and promotion was also a practical issue which was highlighted, concerns were expressed that content was not actually reaching the wider community because they are unaware and thus content just gets shared between community reporters. Another practical measure which is quite key is the setting up of a Salford reporter's mail out system where the organisation sends out information about events and assignments which people can volunteer to cover. Community reporters also use this system to contact each other.

The symbolic dimensions of domestication relate to the way social actors interpret and give the technology meaning. People's Voice Media are essentially giving social media a very particular meaning from the outset in terms of 'community reporting' and empowerment which they see as people telling stories about their lives or where they live and putting them online. An investigation of the motivations of community reporters reveals that while they may have specific personal goals relating to social or economic development, many share this interpretation of the value of social media. Thus in essence community reporters and those who support it are a 'community' in themselves in that they share similar values. Community reporters, for example, explain that they want to 'give a voice to the voiceless', 'to empower disabled communities', to give a voice to the people of Salford and Manchester.' They also have a shared identity symbolised by the organisation in the form of a community reporter's badge, which some see as empowering in itself in that, for example, they can get into events for free and so on. However other social actors revealed that social media can also be interpreted as a government driven method of control, reflecting the contingent nature of domestication. On the one hand, for example, funding in Salford has been put into social media linked to notions of empowerment; on the other hand the local *Salford Star* magazine, which already provided a voice on regeneration issues, has lost its public funding and has been forced onto the Internet. This has actually had an adverse effect in terms of distribution within the local community limiting access to those who do not have access to the Internet which was identified by one respondent as still a major issue within Salford. The question arises the whether 'technological fixes' in a new guise are being applied to areas of deprivation.

Finally in terms of the cognitive dimension, it has long been recognised that effective use of technology requires people to develop the skills to use them, skills thus became integrated within the digital divide discourse (Warchauser, 2002; Gurstein, 2003). People's Voice Media run informal community reporters courses through which participants develop a range of technical skills, from setting up a blog, to using a digital or video camera to editing and uploading content, depending on their interests. They also develop technical and social skills through the practice of community reporting e.g. going out to cover community events, conferences and festivals, which one person described as ongoing training. In addition community reporters get the opportunity to attend an introduction to journalism course at the BBC. Thus, aside from the technical skills this form of social media appropriation demands communication and interpersonal skills i.e. skills such as interviewing people and ability to relate to people. It also demands a certain level of confidence in being able to approach people. Community reporters find it helpful to work in teams because different people have different skills, for example, some excel at interviewing while others excel at using a camera or editing footage. The whole process in which skills are gained is very much about blending the real and the virtual and people need and develop skills which embrace this. For those who have been socially isolated this can be both a challenge and an opportunity. For some like Kevin the mode of training, which is described as a 'throwing you in at the deep end approach,' works well, others felt however that they needed more guidance and structure. Community reporters who took part in the Talk Broughton project, for example, had very specific guidance and support structures in place, while those who work on their own can struggle without clear guidance, for example on what they can or can't produce and ideas for content production. A further issue is where training takes place, for example in the case of Edge Lane Allotments, developing skills on site through bringing a laptop out to the allotments was seen as a vital part in getting people interested in developing technical skills to integrate within the allotment project.

## **Social Media and Digital Inclusion**

This section looks at the implications of the findings for the digital inclusion discourse. Firstly the preceding discussion of social media appropriation within the context of community reporting clearly demonstrates that engaging people in use which is meaningful to them constitutes far more than providing access to technology and skills to use it. Practical, symbolic and cognitive dimensions of domestication all play a role in social media appropriation. Of particular interest is the symbolic dimension of social media appropriation which shapes engagement revealing the relational characteristic of social media technologies (Mork Peterson, 2009). Characteristics of the technology such as mobility, advancing ease of use and pervasiveness of the devices used to capture content also all play a part in engaging people. For example, some community reporters commented that they wanted to get across how easy it is use the technology because some people feel it is beyond their reach or they just aren't capable. Digital exclusion in relation to access to the Internet remains an issue which limits the potential of the content produced. Of key interest in the fact that community reporters are people who would traditionally be configured (particularly within policy related to the digital divide), as socially and digitally excluded generally related to that fact that they may not have a role which is deemed to be economically productive. Community reporters interviewed live in areas which have been socially constructed as deprived and include specific social groups commonly associated with these forms of exclusion; the elderly, asylum seekers, the disabled, the unemployed, the long term sick, people recovering from mental health issues, drug and alcohol abuse. They are in fact being very productive within their communities empowering themselves and others through volunteering as a community reporter. The findings thus challenge key assumptions about people inherent within the digital divide debate and suggest that a focus on specific social groups may be unhelpful. Focusing on the individual and encapsulating diversity in terms of their interpretations of the technology and motivation for use or non-use may be far more helpful approach in understanding why people do or do not engage with social media and whether and how actions need to be taken to promote social media appropriation.

## **Sustainability and free and open source software**

The final section briefly looks at whether the free and open software associated with social media has any implications for the social and economic sustainability issues often associated with community development organisations and community informatics initiatives (Gurstein, 2003). Software such as gimp, wordpress and audacity were identified as helping the organisation financially. Such software has a social sustainability element in that they also perform an important role in that staff can teach people how to use the software and community reporters can go home and download it for free. Audacity was described by one member of staff as a very stable piece of software, the fact that it is cross platform and therefore accessible to people who have Applemacs as well as PCs was also identified as important. The key community sites used by community reporters are also based on the free software called droople. One member of staff said that this makes it sustainable not only because it is free but because the development of the site can be easily be picked up on should the site developer leave the organisation. Some of the free and open software it was felt was not that helpful particularly in relation to video editing and training in the use of final cut software is a particular interest of community reporters in terms of skills development which is expensive but produced a higher quality product. Another issue is that People's Voice Media and therefore the Community Reporters programme is all about accessible technology including in terms of ease of use. One of the staff, David, said that engagement in this context is all about engaging people in something that is easy to use but even though it is free "*some of it can be difficult to get to grips with and open source is quite technical do and is two or three steps down the line.*" The issue of the time involved in figuring out whether it will actually be helpful or not was also raised by David. Overall sustained funding remains an issue for the community development organisation and qualifications and the selling of content have emerged as possible routes for sustainability. People continue to engage with the programme for a variety of reasons some relating to access to and learning skills in technology use such as free software, but for many, other factors such as the social interaction between community reporters keep them involved. Thus it would seem that free and open source software, though helpful, has a marginal role in social and economic sustainability of the initiative.



## Conclusions

The research is still ongoing and there are as yet no clear conclusions from the data examined to date. The analysis utilised within this paper also represents the first stage of data analysis and ideas are being reframed continually in the light of new data and literature. Generally it can be summarised that the case of the Community Reporters Programme illustrates how underlying strategies for domestication and contingent factors both promote and inhibit the potential for community empowerment and regeneration. Particular areas of interest emerging are the impact of the portability of various technologies on social media appropriation and the associated safety and privacy issues. It is also recognised that the discussion within this paper needs to be broken down further in relation to the appropriation of individual artefacts used to capture content. The value of community networks in the age of social networking may be a further area for more in-depth analysis of contributions of social media to community empowerment. The implications of the political push for social media use for other forms of community media as suggested through the case of the *Salford Star* also emerges as a further area of interest. Links between social media, community empowerment and regeneration are evident and being explored through the Community Reporters Programme but it is interesting that some concerns related to regeneration remain unvoiced. Contributions have been made to the digital inclusion discourse, challenging assumptions about people within this and technological deterministic assumptions that the technology is somehow inherently empowering. Open source and free software are found to help the finances of a community development organisation to a degree but not to sustain its activities. Such software appears valuable when it is easy to use, cross platform and when it produces results of a high enough quality. The evidence suggests that sustained engagement is linked to a range of factors and certainly not based purely on the availability of free software.

## References

- Aurigi, A. (2005), 'ICT and the City, Tensions in the Digital City', *The European Journal of Planning*. Available at: <http://www.planum.net/topics.documents/ict-and-the-city.pdf>
- Bijker, W.E., Hughes, T.P., and Finch, T.P., (1987), *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*, MIT Press, Cambridge, MA.
- Boyd, D.M. and Ellison, N.B. (2007), Social Networking Sites History, Definition and Scholarship, *Journal of Computer-Mediated Communication*, vol. 13, no.1. Available at: <http://jcmc.indiana.edu/vol.13/issue1/boyd.ellison.html>.
- Bure, C. (2005), 'Digital Inclusion without Social Inclusion: The consumption of information and communication technologies (ICTs) within homeless subculture in Scotland', *The Journal of Community Informatics*, vol.1, no. 2, 116-133.
- Byrne, D. (2007), 'Public Discourse, Community Concerns and Civic Engagement: Exploring Social Networking Traditions on BlackPlanet.com', *Journal of Computer-Mediated Communication*, vol.13, no.1, article 16. Available at: <http://jcmc.indiana.edu/vol/13/issue1/byrne.html>.
- Clement, A., and Shade, L.R. (2000), 'The access rainbow: conceptualising universal access to information/communications infrastructure', in Gurstein, M. (2000), *Community Informatics: Enabling Communities with Information and Communication Technologies*, Idea Group Publishing.
- Coopman, T.M. (2009), 'Toward a Pervasive Communication environment perspective', *First Monday*, vol.14, no.1, 5th January.
- Cushman, M and McClean, R (2008), 'Exclusion, inclusion and the changing face of information systems research', *Information Technology and People*, vol.21, no 3, 213-221.
- Dabinett, G. (2000), 'Regenerating communities in the UK: getting plugged into the information society'. *Community Development Journal*, April, vol. 35, no.2, 157-166.
- Department for Communities and Local Government, (2008), *Communities in Control: Real People, Real Power*.

- Eubanks, U.E. (2007), 'Trapped in the Digital Divide: The Distributive Paradigm in Community Informatics', *The Journal of Community Informatics*, vol.3, no.2.
- Friedmann, J. (1992), *Empowerment. The Politics of Alternative Development*. Blackwell, Oxford
- Gilmoor, D. (2004), *We the media, grassroots journalism by the people for the people*. O'Reilly Media, Sebastapol, California
- Graham, S. (1997), Urban Planning in the 'information society', *Town and Country Planning*, November, vol.66, no.11, 296-299.
- Graham, S. and Marvin, S. (1999), 'Planning Cybercities? Integrating telecommunications into urban planning', *Town Planning Review*, vol. 70, no.1, 89-112.
- Gurstein, M. (2000), *Community Informatics: Enabling Communities with Information and Communication Technologies*. Idea Group Publishing, London.
- Gurstein, M. (2003), "Effective Use": A Community Informatics strategy beyond the digital divide', *First Monday*, vol. 8, no.12.
- Harris, K. (2000), "Everyone Gets Hooked": exploring ICTs in low-income neighbourhoods, Community Development Foundation, London.
- Hellawell, S. (2001), *Beyond Access ICT and social inclusion*. Fabian Society
- Helsper, E.J. (2008), *Digital Inclusion: An Analysis of Social Disadvantage and the Information Society*. Oxford Internet Institute, Department for Communities and Local Government.
- Hunt, J. and Stout, T. (1997), 'Strike IT Lucky...', *Property Week*, June 27<sup>th</sup>, 34-35.
- INSINC (1997), *The Net Result: Social Inclusion in the Information Society*. The INSINC Report. The Community Development Foundation. Available at: <http://www.uk.ibm.com/comm/community/uk117.html>
- Klecun, E. (2008), 'Bridging Lost Sheep into the fold: questioning the discourse of the digital divide', *Information Technology and People*, vol. 21 (3), 267-282.
- Laverack, G. (2001), 'An identification of the organizational aspects of community empowerment', *Community Development Journal*, vol. 36, no.2, 134-145.
- Laverack, G. and Wallenstein, N. (2001), 'Measuring community empowerment: a fresh look at organizational domains', *Health Promotion International*, vol.16, no.2, 179-185
- Leach, B (2000), 'The DOT, Community Models for Local Employment and Enterprise', paper for *Global CN 2000*, Barcelona, November 2-4, 2000. Available at: [www.mmu.ac.uk/h~ss/sis/bi/barca1.htm](http://www.mmu.ac.uk/h~ss/sis/bi/barca1.htm)
- Li, C and Bern off, J. (2008), *Groundswell: Winning in a world transformed by social networks*, Harvard Business School Press, Boston.
- Loader, B.D. (2002), 'Whither Electronic Communities? Community Informatics in Principle and in Practice', A paper for *Community and Information Technology: the Big Questions*, 16<sup>th</sup> October, Melbourne, Australia: Monash University. Available at: [www.ccn.net/searchcont/loader.htm](http://www.ccn.net/searchcont/loader.htm)
- Loader, B.D. and Keeble, K. (2004), *Challenging the digital divide? A literature review of community informatics initiatives*. The Joseph Rowntree Foundation, York.
- Miller, D. (1999), Community Information Networks. In Pantry (ed.,) *Building Community Information Networks: strategies and experiences*. Library Association Publishing, London.
- Myers, M.D. (2003) (ed.), *Qualitative Research in Information Systems*. Available at: <http://www.qual.auckland.ac.nz>
- Page, N. and Czuba, C.E. (1999), Empowerment: What Is It?. *Journal of Extension*, vol.37, no.5. Available at: [www.joe.org/joe/1999october/comm1.html](http://www.joe.org/joe/1999october/comm1.html)
- PAT 15 (2000), *Closing the Digital Divide: information and communication technologies in deprived areas*. Department of Trade and Industry, available at <http://www.pat15.org.uk>

- Peterson, S.M. (2008), 'Loser Generated Content: From Participation to Exploitation', *First Monday*, March, vol. 13, no. 3.
- Potter, A.B. (2006), 'Zones of Silence: A framework beyond the digital divide', *First Monday*, No.5, May.
- Sassi, S. (2005), Cultural differentiation or social segregation: Four approaches to the digital divide, *New Media and Society*, vol.17, no.5, 684-700.
- Selwyn (2004), 'Reconsidering political and popular understandings of the digital divide', *New Media and Society*, vol. 6, no.3, 341-362
- Shrock, A. (2009), 'Examining social media usage: Technology clusters and social network site membership', *First Monday*, January, vol.14, no.1,
- Sherman, C. (1999), 'The Third System, New Technologies and Employment. An analysis based on the experiences of TSEP and the Article 6 projects'. Extract from paper prepared for the *Third System and Employment Capitalisation Project*.
- Silverstone, R., Hirsch, E., and Morley, D. (1992), 'Information and communication technologies and the moral economy of the household', in R. Silverstone and E. Hirsch eds., *Consuming technologies: Media and information in domestic spaces*, Routledge, London, 15-31
- Sørensen, K. (2005), 'Domestication: The enactment of technology', T. Berker, M.Hartmann, Y. Punie, and K. Ward eds., *Domestication of media and technologies*, Open University Press, Maidenhead, 40-61.
- Southern, A. (2002), 'Can information and communication technologies support regeneration?' *Regional Studies*, vol.36, no. 2, 697-702.
- Southern, A. And Townsend, A. (2005), 'Information and Communication Technologies and their role in Urban Regeneration', *Local Economy*, vol. 20, no.3, 266-279.
- Stoecker, R. (2005), 'Community Informatics good for communities? Questions confronting an emerging field', *Journal of Community Informatics*. Available at: <http://www.ci-journal.net>
- Van Lieshout, M., Egyedi, T. M., & Bijker, W. E. (2001), *Social Learning Technologies: The introduction of multimedia in education*, Ashgate, Aldershot.
- Warschauer, M. (2002), 'Reconceptualising the Digital Divide', *First Monday*, July, vol.7, no.1
- Warschauer, M. (2004), *Technology and social inclusion, rethinking the digital divide*, MIT Press.
- Wilson, P. A. (1996), 'Empowerment: Community Economic Development from the Inside Out', *Urban Studies*, vol. 33, no.4, 617-630.
- Wyatt, S. (2005), 'Non-Users Also Matter: The Co-construction of Users and Non-Users of the Internet' in Oudshoorn, N. and Pinch, T. eds. *How Users Matter: The Co-construction of Users and Technology*, MIT Press.
- Zheng, Y. and Walsham, G. (2008), 'Inequality of what? Social exclusion in the e-society as capability deprivation', *Information Technology and People*, vol.21, no. 3, 222-243.

## An Introduction to the Poetry of Selima Hill

**Lucy Winrow**, *School of English, Sociology, Politics and Contemporary History, University of Salford*

Selima Hill has been called "one of the least embarrassed poets to have found publication" (O'Brien, 1998: 257). In this paper, I will examine this statement in relation to how Hill constructs gender in her poetry. To do this, I will focus on the themes of mental illness from the perspective of the female patient and the mother/daughter relationship.

Judith Butler's work on gender in *Gender Trouble* and *Bodies That Matter* will provide the background to my understanding of gender identity, alongside some sociology-based theories of embarrassment. Simone De Beauvoir's statement "One is not born a woman, but rather becomes one" (qtd in Butler, 2006:1) is central to Butler's argument. Butler argues De Beauvoir "meant merely to suggest that the category of women is a variable cultural accomplishment...and that no one is born with a gender" (2006: 151). Butler asserts that there are only two forms of acceptable gender within society, male and female, resulting in a binary relationship. She explains that an oppressive regime of exclusion has been vital for the binary to exist and emerge as necessary:

Gender is the repeated stylisation of the body, a set of repeated acts within a highly rigid regulatory frame that congeal over time to produce the appearance of substance, of a natural sort of being. (2006: 45)

This 'agreement to perform' on the basis of exclusion of those identities that do not fit results in punishment for those who fail to comply. Those who fail to behave as society expects them to in terms of gender can face discrimination and repression. In *Gender Trouble*, Butler proposes an alternative to the problematic way in which the binary produces and organises identities:

This radical formulation of the sex/gender distinction suggests that sexed bodies can be the occasion for a number of different genders, and further, that gender itself need not be restricted to the usual two. (2006: 152)

It may be that by changing or stopping these performances which are attributed to gender, giving it a naturalised appearance, that a new and more encompassing approach could be developed: "The possibilities for gender transformation are to be found precisely in the arbitrary relation between such acts, in the possibility of a failure to repeat" (2006: 192). A crucial question then emerges, "what kind of subversive repetition might call into question the regulatory practice of identity itself?" (2006: 44). It will be argued that a possible answer lies in Hill's construction of gender identity, and its subversive repetition within her poetry.

Embarrassment is relevant to Hill's work in two ways: firstly in the formation of gender identities and secondly the use of *intentional embarrassment*, as a form of challenge to the reader. John Money's work on intersex helps to illuminate the former point. The intersex condition parallels some of the struggles experienced by women as a result of the strict heterosexual binary. Money believes that there is 'sexual neutrality' at birth which only becomes differentiated as masculine or feminine over time, as a result of experiences growing up (Kitzinger, 2004: 451). Following this initial period of gender plasticity, it could be that experiences of embarrassment play an important role in eventually 'fixing' gender. In developing these ideas, sociologist William F. Sharkey defines embarrassment as being, by necessity, something social in that other people are present. Also, that it can be the "subjective experience of anxiety or fear that one's behaviours will be negatively sanctioned or others will have a lower evaluation" of the individual concerned, similar to the way in which 'inappropriate' behaviour and identities are punished within the rules of the heterosexual binary (Kalbfleisch, 1993: 147). Finally, Sharkey argues that a discrepancy must exist "between the person's presented self and one's desired presentational self" (Kalbfleisch, 1993: 147).

It is this 'gap' that may account for some of the tension within Hill's work in relation to gender. Sharkey's work on intentional embarrassment supports the argument that for Hill, this is a

highly effective method of challenging beliefs about gender roles. Sharkey explains that intentional embarrassment can be used for the purposes of "transmit(ing) a message for eventual interpretation by a receiver" (Kalbfleisch, 1993: 148). This will be important when considering Hill's work in terms of the message that is intended, and finding out who could be embarrassed by her work and why. Whilst embarrassment is perceived to be an anti-social behaviour because it violates taken-for-granted rules of interaction, Sharkey maintains that violations of expectations may be allowable as long as positive valued actions accompany or follow the violation, thereby compensating for the violation. He concludes that the intentional embarrassment of another can "cause a momentary degree of uncertainty that may prove to have a positive relational outcome" (Kalbfleisch, 1993: 157). Could it be that through embarrassing her reader, Hill destabilises their beliefs enough in order to chip away at the constraints on the body, the highly ordered rules through which society sees gender?

Hill's personal experiences seem to have influenced the versions of gender identity which appear in her work. She describes grappling with her own identity whilst growing up and the idea of a "women's language" (*Poetry Quartets 2*, 1998). Her strict education and the great influence of her father, led her to feel as if she had a "man's brain in a woman's body" (*Poetry Quartets 2*) and furthermore, she was aware of being able to impress male editors with her writing. It was only later, and with the help of other female artists, that she felt able to write "from a woman's body and from the heart" (*Poetry Quartets 2*). Deryn Rees-Jones addresses this remark:

Her [Hill's] sense of herself as a poet thus depends upon a gendering which draws on stereotypical divisions between masculinity and femininity and the setting up of a gendered conflict between the mind and the body. Such a construction perhaps feeds into her acute anxieties about the relationship between the construction of her poetic 'I' and its relation to her own bodily self. (Rees-Jones, 2005: 165)

This clear source of tension within Hill's work will be demonstrated later in a discussion of the poems. In discussing her methods as a writer and how she finds her 'voice' not just as an artist but as a woman, Hill describes attempting to write in the first person and finding it difficult to proceed. Things became easier however when she changed from 'I' to 'she', abandoning "the little island of me" (Herbert & Hollis, 2000: 236). This may suggest that there is a difficulty in writing from the gendered self because the self is hard to define; there also seems to be a constant pressure to live as a particular gender which is not easily forgotten, even when writing. However, whilst this proves for Hill to be a constraining condition, it may be that others could feel more comfortable writing from a prescribed, clearer position that fits into the social conditions of male or female.

Hill approached her collection *Violet* (1997) by adopting a number of narrators such as the ex-husband, a nurse and a patient, in order to explore the bodily self and perhaps attempt to resolve or explore the tensions in herself as a writer. This was perhaps an easier way of negotiating gender identity; to show that the female character need not be limited to stereotypically 'female' behaviours but could adopt a number of different voices.

Having been a mental health patient herself, as well as volunteering in prisons and hospitals, these experiences are apparent in Hill's writing. The focus of her work is the struggle faced by women with how they are expected to behave. *The Accumulation of Small Acts of Kindness* (1994) is based on Hill's experiences of mental breakdown, as well as time spent as a volunteer in the Maudsley Hospital in South London. We are told in the foreword that the book is written "in code", with voices "real and imagined, read or remembered." It may be that the protagonist fears exposing her true self, along with confusion and guilt as to who she is or is allowed to be. The text is littered with female body parts, such as the breasts, uterus, vagina, as well as female conditions such as pregnancy and menstruation: this will be illustrated in the later discussion of the poetry. There is a great deal of anxiety and anger attached to these parts and feelings of disconnection. Despite this, Hill makes a determined attempt not to ignore the woman's body but to interweave it with the protagonist's experiences. In terms of embarrassment, there is a room in the hospital in *The Accumulation* "where the good girls go" (1994: 77) which implies that there are bad girls and that behaviour is rewarded or punished accordingly. In Sharkey's theories, behaviour that is not socially acceptable (in this case, it is the behaviour of a gender which is inappropriate) is negatively sanctioned. This suggests that in order to avoid confrontation it is perhaps easier for some bodies to perpetuate their current

gender roles given that the protagonist is chastised for failing to uphold her socially prescribed gender.

A prominent source of anxiety for Hill's characters arises from approaching womanhood, sex and men. In Chapter 1 of *The Accumulation 'Masters'*, the protagonist is described as "departing" for womanhood (77). Part way into the book we discover she may be pregnant: "Being pregnant feels dishonest. / Now I sleep with foreign men in dreams" (76). She clearly feels shame about her body and what it is physically capable of, as well as sadness about the loss of her freedom which is now confined to dreams. The figure of the virgin appears throughout the book, often depicted as admirable "Taut purity of virgins in white landscapes" (59). However, it has sinister and uncomfortable overtones, such as the "virgin in her wooden dress" (73), whose clothing preserves modesty and makes her inaccessible yet also hints at death, through its coffin-like appearance, literally imprisoned by her gender. Hill is questioning this construction through the tension it causes within her characters. This hostility felt towards pregnancy and femininity may be another cause of difficulty for the protagonist who makes frequent reference to her own writing. There is a suggestion that a woman cannot be both a mother and an artist, "whenever a child is born, a woman is wasted / we do not quite belong to ourselves" (68). To this the protagonist adds, "the writer's instinct is essentially heartless". The implication is that she is being incarcerated for holding these views, as if a 'normal' woman would not say such things. In everyday life, it would be embarrassing to admit such feelings.

Even for those women who abandon all notions of femininity, this does not necessarily bring freedom. In Hill's collection *Lou-Lou* (2004), the psychiatric patients explain "We've forgotten our houses, our homes / our pretty clothes / our little dogs on leads / we're not that kind of women anymore" (59). These women are wholly un-feminine and are depicted as unattractive, dribbling and insane: something which is uncomfortable and embarrassing for the reader to witness. It is as if that is the only alternative to pretty dresses, operating as a stark warning about what happens if the social rules regarding gender are not upheld. Sadly, the implication seems to be that the treatment the women are receiving is not being offered to help and heal but to make them fit back into that role again, "This is kindness / Leading me back" (64), i.e. not forward.

The ward is filled with many strange and unfamiliar images of 'failed' womanhood, from the grinning, serene woman who "apparently killed a baby once" (57) to the chilling poem about the "tiny married women" (56) whose husbands put them there. Regardless of the story, patients greet new arrivals with a similar level of tolerance, "They've got to learn to love it here, / like we do." (56) When dealing with the patients directly, the doctors make them feel very uncomfortable, brutishly pressing against them like "bristly pigs...but less discreetly" (51). They are also depicted as hardened to the pain and suffering of the mentally ill, indifferent to it,

The doctor disappears  
like a goods train  
that travels through the night without stopping  
or caring who is born or who dies. (43)

Much of what goes on is perhaps not apparent to a casual visitor, covered by the nurses' sugary tones, "Look, she says, / aren't they beautiful?" However the patients are intensely aware that they would "probably want to strangle" their "beautiful necks" (44). Despite all the treatments, there is still a lack of understanding about mental illness, for all involved. The narrator asks us "Who but a beaten specialist would offer/ to map the changes of the ocean floor?" (80). It is as if the experts are in unknown territories and unable to fully understand the ever changing mental states of the patients. There is fear of the stigma attached to the waiting room with its "diseases of the mind." It may be an embarrassing experience for the reader to discover that they had held similarly prejudiced views towards the mentally ill and women who did not conform to the expectations of their assigned gender. The protagonist in *The Accumulation* knows that talking isn't necessarily communicating, "I haven't talked to anyone for years" (82). If anything, it is to be avoided "Someone to talk to, cold as charity" (78).

Seemingly no longer capable of existing in the outside world, maternal instincts are replaced by thoughts of medication and abandoned children:

We creep about the ward at night  
with sleeping pills nestling  
in the pockets of our dressing-gowns  
like the motherless eggs  
of a sort of mutated insect  
that doesn't know or care about tomorrow. (18)

As well as being distressing for the patients, Hill addresses the effect it has on the family, who often find the situation hard to understand. The patients are depicted particularly unpleasantly during visiting times

Our visitors go home  
as soon as possible  
but even there  
they hear our low grunts. (28)

The visitors feel guilt at the shame and embarrassment they feel, seeing their animal-like, snorting relatives. The patients are even described as fat frogs "squatting on the feet of their beloved" (48) evocative of the burden they present. When the protagonist eventually reaches a stage where she is deemed ready to leave, she observes the family and friends of other patients who they "hope and pray will one day be normal" (63). There is still no explanation of what normal is, just the worrying thought that the only thing on the protagonist's mind is to walk out slowly as "hasty movement" (63) doesn't suit her hair. The lines: "They lead us away to be normal / hair-dos swaying" (64) seems as if the protagonist is robotic, brainwashed to act accordingly with her gender with the implication that nothing has really changed and everything will unravel again in the future. As has been shown, Hill does reproduce some stereotypes in her depictions of the mentally ill with their animal-like behaviour and noises evoking fear and repulsion, along with examples of the misunderstanding that surrounds mental illness. She does however challenge these perceptions within her work by presenting characters whose anxieties are commonplace in everyday life but are perhaps not expressed. There is a distinctly human element in their vulnerabilities, heightening the feelings of alarm as to how these actions are perceived when displayed in a female and the context of a hospital. The reader may feel that this frightening place is less distant to them than they once thought. This in itself can provide valuable insight into current attitudes to gender and the behaviour which is categorised socially as 'normal'.

Another area in Hill's work where the construction of female gender is explored is in her depiction of the mother/daughter relationship. As discussed, for some of Hill's characters, the expected maternal feelings are not present and instead there is the desire to do more with their lives than just bear children. They seem to resent their bodies for having the capacity to give birth and to define them. This extends to the feelings towards the born child. In Hill's collection *Bunny* (2001) we learn of the resentment that exists between mother and daughter after the daughter is burnt in a house fire as a baby. Hill herself was burnt in a fire as a baby and whilst it cannot be clear what her feelings about that are, the protagonist in the book seems to blame her mother whilst at the same time feeling that she herself is inadequate and ugly. This is revisited in Hill's 2008 collection *The Hat*. 'The Day the Cows Got Out' (2008: 9) evokes the panic of a fire on a farmyard with parents referred to rather impersonally as "the mother, the father." Amongst the drama, in which the cows are loose, the mother figure is seen as small and insignificant, getting in the way of the firemen. The protagonist feels such guilt and self loathing that she declares that there "shouldn't have been a baby in the first place" (9). From almost the moment of birth, this individual is marked as wrong and shameful. It is not clear if this is the protagonist's voice or the repetition of her mother's beliefs, most likely the two are linked. There is an accusatory tone towards the mother,

her clothes are stuck to her skin, I'm afraid,  
and all the tweezers in the world, Mother,  
will never get them tweezered out again.

There is also the implication that this now damaged child was passed on to a wet nurse, rejected by her mother, "easy does it, nurse-, easy does it. / Drip that milk." The exclamation "Bastard!" that follows carries great force and sets the tone for a dysfunctional

mother/daughter relationship throughout the rest of the collection, as well as perhaps making a statement about the child's legitimacy.

Indeed as the girl begins to reach maturity, the tension between herself and her mother begins to escalate. The poem 'Gentle Pressure on the Sides of Piglets' (2008: 11) introduces a daughter who is trying to calm her mother. The mother is depicted as a piglet, seemingly sweet and non-threatening, however we learn that she "crushes underfoot" the snails that she is examining. Perhaps she is too clumsy and thoughtless in her approach, mirroring how she treats her daughter. 'The Darkness of Her Meekness' (21) shows how the mother can be complicit in keeping her daughter as a slave to her gender. Here, the mother has "installed her as in stalls" (21) into this darkness, as if this is where she belongs and needs reminding of that fact. While this is happening we are told "in the lounge her mother and her sisters / further their spectacular depressions." It is almost as if these characters are wallowing in the pain of what their role as women requires of them, martyring themselves and the protagonist must too. 'The Sort of Neck People Want to Strangle' is a sinister poem with the protagonist's neck being described as snowy, like a goose that is "sweet but nasty" (24). It may be that she is nasty because she is sweet? Is it because she is a woman or because she is too pretty? Chillingly the mother, traditionally a caring role, replies "yes she can" to the question of whether she can imagine strangling her daughter. The reasons behind this are complex. She may be jealous of her daughter's youth and appeal to men or angered by her daughter's stance against that of the stereotypical female role. She may feel guilt over the house fire and her own failings as a mother.

The poem 'Bicycle' reinforces the suspicion that the protagonist's mother believes she needs a man to put her in her place: "Her mother's found her someone she can *lean on* - / as if she were a bicycle!" (2008: 28). Being depicted as a mute, inanimate object, in this case, something which can be ridden, makes her inherently submissive. The protagonist imagines being a "sparkly married bicycle / which wouldn't dream of being mainly throat; / which wouldn't dream of being eased open." This takes the idea of being an object one stage further to the point that her own body parts don't belong to her. The implication is that she is quietly and without making a fuss, to be used for the sexual pleasure of her husband. In 'Violence' "her heart is like a garden full of wounds." The pain that she is experiencing is associated with biology and childbirth. It aches like "mothers / whose stiff nipples/ache for mouth, / and cast around for them violently" (44). As with the protagonist's mother, the women here have very little maternal 'instinct' in a caring way, theirs is more desperate and alarming. It is as if they are nothing without a baby and their body is useless if not for child bearing. To describe the protagonist's pain in this way is to suggest this is the only pain women feel, or are allowed to feel.

Hill wrote an article for *Contemporary Women Poets: Reading Writing Practice* (2000) called 'God's Velvet Cushions' in which she states that colour is one of the defining aspects of her work and explains that it is integral to the actual writing process from the beginning, "Poems for me are colours, they begin, are grounded in, informed by, unified by colours and their relationship of colours" (Mark & Rees-Jones, 2000: 29). In *Bunny*, the colour blue binds together the fragmented poems, as well as representing different moods with different shades. Often used to depict the foreboding night time, the colour also suggests hope as in the midst of her troubles, the protagonist senses there may be a way out whilst studying an 'Atlas of the Heavens' "in whose blue circles everything is possible / by virtue of the beauty of their names" (2001: 31). Additionally, in terms of a representation of female experience, the colours red and white appear frequently throughout Hill's work. In *The Accumulation*, red is seen in menstrual blood, miscarriage and violence, whereas white represents milk and the purity of the virgin. There seems to be a conscious effort to confront these uncomfortable topics. The female body is presented as hostile and threatening, "My breasts are rocks of milk. / They found the razor" (1994: 66). Coupled with "My uterus is like a sunlit knife" (68) there is implied red of bloodshed and intentional harm, both to the child but also possibly of self harm. At times we are told "My socks are blood-soaked" (61) and "I think I'm bleeding" (64) suggesting either menstruation or miscarriage and evoking the uncontrollable nature of the female body and why it fails to fit neatly into the binary. Perhaps in an attempt to represent the horror that is felt towards female bodily functions, Hill places alongside these descriptions food that is either red or white. A baby is described as like a plum being fished out of custard and falling asleep "with prune juice in their hair" (67) which creates the unpleasant image of consuming these bodily fluids. The colour white is also used to refer to the inadequate and impersonal treatment offered to mental patients, the protagonist suggests "The best thing



they could do is to invent / a nice white dye that they could dye us all with." (61). The desire seems to be for them to be all the same and 'pure', to quietly be put in their place as well as the image of killing with the pun on 'die'. This takes a foreboding turn when shortly after this she talks of being given white grapes for supper, as if the food is being poisoned.

This colour symbolism continues in *The Hat*. 'Lily and Beef' begins with the jokey tone "Funny isn't he" but the reality is disturbing as we learn that 'he' is deliberately burning one of the protagonist's breasts, leaving one red and one white, hence the title (2008: 19). There are sexual overtones to burning this part of the body, perhaps for his pleasure or as a slight to the female form. There are also links back to the burnt baby, the re-opening of old wounds. The protagonist seems embarrassed and tries to cover up what is happening "this is just between ourselves ok" (91). This could be because she feels that her role as female in society is to keep quiet and accept how she is treated, she does not want to be seen making a fuss. The embarrassment, and the feeling that what is happening might be wrong, is something that comes from saying it aloud: this is fundamental to Hill's own writing, the speaking aloud of things that are usually hidden, almost as if nothing is said then nothing is wrong. In many of the poems, Hill's characters are transformed into a goose against which there are frequently violent references to strangling or eating, evoking the image of blood on white. Many titles evoke this including Sugar, Semolina, Snow and Goose and the implied red of the violence to be done to the body. In 'Tomato Dreamer' she describes herself as a "furious little strawberry" sinking into a "snowdrop wood" of her love for a man (49). This image shows determination to escape or get to the bottom of something that is completely opposite to her and is swallowing or smothering her. Love therefore seems dangerous and threatening.

The themes discussed of mental illness and the mother, as well as the use of colour have shown how Hill is attempting to write about those elements of female experience which in everyday life, due to embarrassment, are not often discussed. In her depiction of mental illness, Hill is very effective in highlighting the 'gap', which is said to exist in embarrassment theories, between the presented self and the desired presented self. Hill clearly identifies this as a source of tension, showing the way in which her characters think and feel contrasted with how they are perceived by the doctors, family and even the reader. However, by writing about subjects and feelings which are seldom expressed for fear of embarrassment, Hill is intentionally embarrassing her reader who must confront this feeling if they are to continue reading the book. Although this may make them feel uncomfortable, it can help them understand how they themselves have been shaped as a certain gender and the pressures that exist.

In the mother/daughter relationship, the 'growing up experiences' discussed earlier, which it was argued helped to shape gender in the infant, are evoked. From a young age, Hill's female characters are bombarded with images of how women are expected to present themselves, as well as the repercussions within society if that behaviour is not upheld. For some characters, the meaning placed on the bodies into which they were born is so burdensome that they begin to hate those bodies: something which is uncomfortable to admit and talk about. Hill uses stereotypes of mental illness and of traditional women's roles, but then seeks to show the human stories underneath that can be hidden, yet evoke empathy and recognition within the reader. The reader may begin to question their own attitudes and unearth the constructed nature of behaviour and the expectations that they and society place upon themselves and others.

## References

- Butler, J (1993), *Bodies That Matter*, New York: Routledge
- Butler, J (2006, 2<sup>nd</sup> edition), *Gender Trouble*, New York: Routledge Classics
- Herbert, W. N. and Hollis, M. (eds) (2000), *Strong Words: Modern Poets on Modern Poetry*, Newcastle: Bloodaxe Books
- Hill, S (2001), *Bunny*, Newcastle: Bloodaxe Books
- Hill, S (2008), *The Hat*, Newcastle: Bloodaxe Books
- Hill, S (1994), *Trembling Hearts In The Bodies Of Dogs*, Newcastle: Bloodaxe Books
- Hill, S (1997), *Violet*, Newcastle: Bloodaxe Books
- Hill, S, Duffy, C A, Rumens, C, Adcock, F (1998), *The Poetry Quartets 2: v 2* [Audio Book], Newcastle: Bloodaxe Books
- Kalbfleisch, P. J. (1993), *Interpersonal Communication*, New Jersey: Lawrence Erlbaum Associates
- Kitzinger, C (2004), 'The Myth Of The Two Biological Sexes', *The Psychologist* 17 (8)
- Mark, A, and Rees-Jones, D (2000), *Contemporary women's poetry: reading/writing/practice*, Hampshire: St. Martin's Press
- O'Brien, Sean (1998), *The Deregulated Muse: Essays On Contemporary British & Irish Poetry*, Newcastle: Bloodaxe Books
- Rees-Jones, Deryn (2005), *Consorting with Angels: Essays on Modern Women Poets*, Newcastle: Bloodaxe Books