STRUCTURING THE UNSTRUCTURED DATA: THE USE OF CONTENT ANALYSIS

U. Kulatunga, D. Amaratunga, R. Haigh

Institute of Built and Human Environment, University of Salford, Salford, M5 4WT, UK.
E-mail: U.Kulatunga@salford.ac.uk

ABSTRACT: Content analysis is a research technique used to organise large amounts of textual data into standardised formats which allows arriving at suggestions/conclusions. Content analysis can be carried out quantitatively by counting the words or qualitatively by coding. The former approach refers to counting the frequency of the keywords and the later refers to identifying similar themes or concepts from the data set. This paper discusses the use of conceptual content analysis by using computerised software to analyse data gathered from semi-structured interviews. The context of the research within which content analysis is used is to identify the influence of performance measurement towards construction research activities. The paper first explains the research methodology pertaining to this study by reasoning out the selection of case study research approach coupled with semi-structured interviews. The paper then discusses how the information gathered from semi-structured interviews is fed into the computerised software to identify and generate main concepts of the study.

Keywords - Case study, Content Analysis, Research and Development, Semi Structured Interviews.

1   INTRODUCTION

The data analysis of a research project is one of the significant parts of any research as it helps to investigate the collected data and to draw up conclusions based on them. According to Jorgenson (1989, p: 107) data analysis starts with "…breaking up, separating, or disassembling of research materials into pieces, parts, elements, or units". Thereafter, the researcher sorts them and looks for types, sequences, patterns or finds evidence to address the initial propositions of the study (Yin, 2003). The aim of this whole process is to assemble or reconstruct the data in a meaningful way (Jorgenson, 1989). As stated by Hartley (2004) data analysis helps to generate theories which are grounded in the empirical evidence.

This paper discusses the data analysis process carried out in a pilot case study to explore the influence of Performance Measurement (PM) towards construction Research and Develop (R&D) activities. First, the paper briefly explains the background to the study by highlighting the importance of PM to construction R&D activities. This is followed by the discussion on the selection of case study research approach. Next, the use of content analysis and how the categories and codes are worked out for the study is discussed. Further, the possibilities of using computer aided software to assist the data handling process of content analysis are also discussed. The paper is concluded with an explanation of how the content analysis helps to *structure* the *unstructured* data

1.1   Performance measurement of construction R&D

The success of construction research activities is governed by the scientific and technical capabilities of the research providers as well as on its links between research users and beneficiaries (Cripps et al, 2004; Fairclough, 2002). Yet, a lack of clear objectives and lack of successful partnerships between construction research funders, research providers and research beneficiaries is evident (Print, 1999). Even though the importance of understanding

the target audience of R&D activities are well highlighted (Flanagan and Jewell, 2006), construction research is being often accused of not addressing the real life organisational problems, and for being of low level relevance/applicability to the industry needs (Barrett and Barrett, 2003; Townsend, 1999; CRISP consultancy commission, 1999).

Courtney (1999) argues that R&D returns should be "more calculable" by means of establishing certain and visible relationships between the investments and output of the R&D activities. Thus, Print (1999) suggests the implementation of PM, within construction R&D activities, to evaluate the successfulness of research projects. Such a system will ensure the formulation of clear aims and objectives by identifying the scope of the project and the potential research markets/ research beneficiaries. This will, in turn, increase the "take-up" rate of the research outcome and the exploitation of the research results. Thus, the use of appropriate performance measures will safeguard the expectations of the parties involved within the research process and will build up a good relationship between research beneficiaries, research providers and funders. This indicates the need forf Performance Measurement within the construction R&D activities (see Kulatunga et al, 2006a). Accordingly, this study aims to evaluate the influence of PM towards construction R&D activities. The following research questions are formulated with the intention of addressing the objectives of the study:

- What is the current status of construction R&D activities?
- What are the critical success factors of construction R&D activities?
- How are the performance of construction R&D activities measured?
- How can PM influence the effectiveness and efficiency of construction R&D activities?

To investigate this, the case study research approach is selected which will be discussed in the following section.

2   CASE STUDIES AS THE RESEARCH APPROACH

As per the aim and research questions stated above, this study requires exploration of the PM concept within construction research organisations and different views of people identified regarding the critical success factors of R&D, suitable R&D performance measures and measurement techniques and the impact of PM towards construction R&D. Thus, a research approach which facilitates an in depth analysis and the gathering of professional opinions, are required. According to Zonabend (1992) case studies are carried out in a way that incorporates the views of the "actors" of the case under observation. Due to the open ended inquiry used in case studies, it is suitable to build theory and generate hypothesis (Amaratunga et al, 2002). A similar comment is given by Patton and Appelbaum (2003, p: 67), who identify the ultimate goal of case studies as "to uncover patterns, determine meanings, construct conclusions and build theory". Therefore, case studies are selected as the most suitable research approach for the study over other research approaches such as Experiments, Surveys, Action Research, Ethnographical studies (see Kulatunga et al, 2006b). The following characteristics could be listed as the key points behind the selection of case study research approach for this study:

- facilitates an in depth study of the selected environment;
- allows multiple sources of evidence to increase the validity of the collected data;
- does not control/ manipulate the environment under examination (as in the case of experiments and surveys);
- does not interfere with the attitudes, perceptions or the procedures of the environment (as in the case of action research);

- analyses a contemporary event.

Multiple case studies will be selected, as the evidence gathered from multiple cases is often considered more compelling and leads to a healthier and stronger study (Herriott and Firestone, 1983). Each case will be carefully selected to predict the similar results, or to produce contrasting results for predictable reasons (Yin, 2003).

Case studies provide the opportunity to deal with a full variety of evidence (Yin, 2003). For this study, semi structured interviews were selected as the main data collection technique supported with document reviews and a questionnaire survey within case study organizations. Case study protocols, consistent interview guidelines, preparation of case study database, etc. were designed to maintain the reliability and validity of the results. Table 1 shows the case study research approach pertaining to this study.

Table 1: Stages of the case study

Stage	Analysing Mode	Objectives
Stage 1	Semi structured interviews Document reviews	• Identify the background information about R&D process • Identify the success factors of R&D activities • Identify the issues related to R&D management and evaluation of its performance • Identify the expected benefits of performance measurement of R&D activities
Stage 2	Reflection and analysis of data from Stage 1	• Evaluate the strengths and weaknesses of R&D management, its performance evaluation identified from Stage 1 • Group the success factors and development of *critical* success factors of construction R&D
Stage 3	Questionnaire survey	• Identify performance measures for critical success factors • Development of the R&D PM framework

As explained above the primary aim of this study is to explore the PM concept within the construction R&D activities. As shown in Table 1, it is necessary to gather different views regarding the current issues relating to performance evaluation of construction R&D and to identify the perceptions of respondents about the success factors of construction R&D activities. Moreover, it is necessary to evaluate the degree of importance the respondents are attaching to PM in construction R&D activities. Therefore, the data collection medium should facilitate the gathering of differing views and opinions of respondents. As noted by Silverman (2001, p: 87) the interviews in social science strive "…to generate data which gives an authentic insight into people's experience". To facilitate these requirements, the researchers wanted to select an open ended nature of inquiry. However, the researchers also devised a structure/direction to prevent the respondents deviating from the scope of the study and to keep them focused towards the main issues. Thus, semi structured interviews were identified as the most suitable data collection method.

Due to the open ended nature of inquiry, the data gathered was of an "unstructured" nature consisting of long paragraphs, similar concepts in different locations of the text and with unrelated data to the study. Further, as described by one of the respondents of the study "in a semi structured interview the interviewee does not *behave properly* and starts answering questions which will come later on". As a result, the answers for a particular question could

be mixed up with another question. Thus, the researchers had the challenge of "structuring" these "unstructured" data into meaningful and analyzable units to draw up conclusions as the study progresses. The sections below describe how this challenge was met by using content analysis for this pilot case study.

3   DATA ANALYSIS: THE USE OF CONTENT ANALYSIS

3.1   What is content analysis?

Content analysis can be used to scrutinize written material and recorded communication and thereby used in number of fields such as ethnography and cultural studies, marketing and media studies, in literature and public speaking, sociology and political science. The earlier definitions of content analysis had the component of "quantification" associated with it. For example, according to Kaplan (1943, p: 230, cited in Franzosi, 2004), content analysis "…attempts to characterize the meanings in a given body of discourse in a systematic and quantitative fashion". Berslson (1954) also had a similar view about the quantification aspect of content analysis as he defined it as "… a research technique for the objective, systematic and quantitative description of the manifest content of communication". Within this context, content analysis can be used to quantify the words, concepts or themes and characters in a text. However, in the later definitions, the focus of content analysis has moved on to "inference", "objectivity" and "systematisation" (Franzosi, 2004). Accordingly Holsti (1969, p: 14) defined content analysis as "… a technique for making inferences by objectively and systematically identifying specified characteristics of messages". The Bureau of Justice Assistance (2006) also provides a similar definition by identifying content analysis as a set of procedures for collecting and organizing non-structured information into a standardized format, which facilitates making inferences about the characteristics and meanings of written or recorded material. Krippendorf (2004), one of the seminal authors in the area of content analysis, defines it as a research technique which makes replicable and valid inferences from text to a context of their use.

Content analysis enables the breakdown and organization of large amounts of data into categories or codes (Junginger, 1996; GAO, 1996) which can be words, themes, phrases, concepts or sentences. Such categorization or coding of data makes content analysis a useful tool for data analysis as it facilitates arrival at a manageable and meaningful set of data. Further, content analysis systematically identifies the properties of the categories such as their frequencies of occurrence (Zhang and Kuo, 2001; Wikipedia, 2006). This helps to identify the important concepts from the material.

3.2   Types of content analysis

In the literature various authors categorize content analysis in different ways. According to Krippendorf (2004), content analysis can range from the simplest form of word count to thematic analysis, referential analysis, and prepositional analysis. Colorado State University (2006) classifies content analysis into conceptual and relational content analysis. Authors argue that the above types of content analysis can be positioned within the broad classification of qualitative and quantitative stances as mentioned by Silverman (2001) and Mayring (2000). Some of the main types of content analysis methods are briefly discussed below.

3.2.1.1   Word count

This is also known as textual analysis (Moscarola, 2002), word-based analysis (Jackson & Trochim, 2002) or quantitative content analysis (Silverman, 2001). This approach involves counting the frequency of words in the text. The underlying assumption behind word counting is that the words mentioned most often indicate the important concerns. However, there are several constraints associated with the mere word counting of text. For example, the use of synonyms may underestimate the importance of concepts (Weber, 1990). Furthermore, multiple meanings may mislead the researchers when carrying out word count in content analysis (Stemler, 2001).

3.2.1.2 Conceptual content analysis

In conceptual analysis (sometimes also known as thematic analysis) the text is scrutinized to check the existence and frequency of a concept/theme (Colorado State University, 2006; Krippendorf, 2004). In this method, dominant concepts/themes in the text are categorized into codes (Franzosi, 2004). Instead of counting the frequency of word usage as used in word based content analysis, this approach attempts to find similar cognitions under the same concept (Swan, 1997). Thus, the underlying principle is to identify the occurrence of selected terms within the text. These terms can be implicitly or explicitly related to the concepts/ themes under consideration (Colorado State University, 2006). Even though identifying the explicit terms is straightforward, capturing the implicit terms related to a concept/theme needs to be done with care. As the latter is based on the judgments of the researcher, this can affect the reliability and validity of the data. Thus, the development of a good conceptual/thematic analysis requires the researcher to be familiar with the text and pre-testing of the codes (Franzosi, 2004) to clearly define the implicit terms before starting the data analysis process (Colorado State University, 2006).

3.2.1.3 Relational analysis

This method is also known as semantic analysis (Palmquist et al, 1997). Relational analysis goes beyond the conceptual analysis and examines the relationships between the concepts. The focus of relational analysis, as stated by Colorado State University (2006), is to look for semantic or meaningful relationships. Similar to the conceptual content analysis, relational analysis also begins with identifying the concepts from the text.

3.2.1.4 Referential

Since a "meaning" comprises of other kinds of languages such as background and foreground information, silence and emphasis etc., this method takes into account such factors. Thus, referential content analysis is more suitable to analyse the complexity of language in the production of meaning (Franzosi, 2004).

From this study, the researchers intended to explore the respondent's thoughts about the PM concept within the construction R&D settings. Thus, the opinions and attitudes regarding the PM practices in R&D setting, strengths and weaknesses of the current system were investigated. The respondent's views regarding the successful criterion of construction R&D were also explored. Within this scenario, mere word counting would not lead the researchers to achieve the ultimate goals by deriving major concepts/themes form the study. Furthermore, within the scope of the content analysis, it is not expected to develop relationships among the concepts nor intended to analyse language complexities. Thus, by considering the limitations of the other content analysis types and the requirements of this study, conceptual content analysis was identified as the most suitable method. Having chosen the conceptual content analysis, the next step was to code the text. Different coding types and the coding procedure adopted in the pilot case study will be discussed below.

3.3 Coding in content analysis

Stemler (2001) claims the use of codes and categorization in content analysis makes this tool rich and meaningful. Ryan & Bernard (2000, p: 780) reinforce this statement by saying "coding is the heart and soul of whole text analysis." According to Weber (1990, p: 37), "category is a group of words with similar meanings or connotations". Categories have to be mutually exclusive and exhaustive where former refers when no unit comes in between two data sets and the latter refers when the data represents a comprehensive set of units (GAO, 1996).

One of the core questions raised when dealing with content analysis is regarding the development and definition of categories and codes. According to Krippendorf (2004) "how categories are defined ….. is an art. Little is written about it". Literature reviews, researchers own experiences with the study are good sources to identify concepts and thereby to develop categories and codes. In addition, the text itself can generate concepts, categories and codes. Therefore, codes can be identified before, during and after the data collection (Ryan and Bernards, 2003; Ryan and Bernards, 2000). "Coding is data reduction not proliferation" (Bernard, 2000, p: 446). It is important to have a manageable and a reasonable number of codes for the data analysis process depending on the extent and requirement of the study. While too many codes can make your study cumbersome, too few codes also can produce unreliable and invalid conclusions (Palmquist, 2006). It is recommended to develop the categories and codes closer to the original text by using actual phrases or words in the text (Corbin, 1990). This is known as *in vivo* coding (Bernard, 2000).

There are two main approaches for coding as Inductive and Deductive coding (Krippendorf, 2004; Bernard, 2000; Marying, 2000) sometimes also know as Emergent and Priori coding (Stemler, 2001).

3.3.1.1 Deductive Coding
Deductive coding commences with prior establishment of categories and codes based upon a theory and bringing them in connection with the text (Stemler, 2001; Mayring, 2000; Bernard, 2000). This approach is suitable for the confirmatory stage of a research (Mayring, 2000; Bernard, 2000). The use of pre-established categories and codes leads to the carrying out of a well organized data analysis. Nevertheless, this approach can neglect concepts and categories that do not fall under the pre-established categories.

3.3.1.2 Inductive coding
Inductive coding allows categories and codes to emerge from the text itself. This approach is suitable for the exploratory or discovery phase of research (Mayring, 2000; Bernard, 2000) and used extensively in grounded theory.

However, Miles and Huberman (1994) suggest the use of coding that lies in between deductive and inductive approaches. Accordingly, some categories can be pre-established from the literature and can add more as you go along the text. For the research under consideration, the aforementioned approach was used by pre-defining some of the codes through the theoretical background whilst allowing new codes to emerge from the text itself. The detailed description of coding is given in section 5.

4   USE OF COMPUTER AIDED SOFTWARE

A number of computer software applications have been developed to aid the steps of content analysis which can perform word counting, coding, building conceptual models, linking concepts into networks and producing numerical text by various materials (Bernard, 2000). However, computer aided software needs to be used with caution as they have both strengths

and limitations (Fielding and Lee, 1991). One of the main advantages of computer aided software is its ability to rapidly handle large volumes of data. BUsing computer aided software, text can be easily manipulated and displayed in a number of ways (Robson, 2002). In addition, computer aided software provides a single location to store the materials ranging from interview transcripts, category definitions, interpretations, comments etc. (Robson, 2002; Mayring, 2000). This makes the data analysis process more comprehensive, transparent and replicable thus increasing the reliability and validity of the analysis. On the down side, the use of such software can make the researcher mechanistic and damage creativity.

The authors argue that it is at the researcher's discretion to allow the computers to take the control of the study and to make it mechanistic. If the computer aided software is used with care, to assist the tedious tasks of data handling while keeping the intellectual part with the researcher, such tools can enhance the data analysis process. Accordingly, it was decided to use NVivo (version 2) software for the data analysis process due to the following reasons:

- The number of interviews are fairly high to deal with manually;
- NVivo allows the carrying out of a rigorous and comprehensive data analysis process;
- Increased transparency of the analysis;
- Facilitates use of memos to record important information related to the interviews;
- Easy accessibility to the complete lists of codes which make revisions much easier;
- Codes can be easily retrieved, thus the consistency of coding is maintained while avoiding duplication.

5 DISCUSSION

All the semi structured interviews were tape recorded and transcribed before the data analyzeds. The transcripts in word format were changed to rich text format (RTF) to upload to the NVivosoftware. As mentioned in section 3.2 conceptual content analysis was used for the data analysis process. Accordingly, the transcripts were carefully scrutinised with the aim of identifying concepts related to the research questions. While going through the transcripts, they were broken down into meaningful content categories related to a particular concept (Figure 1).
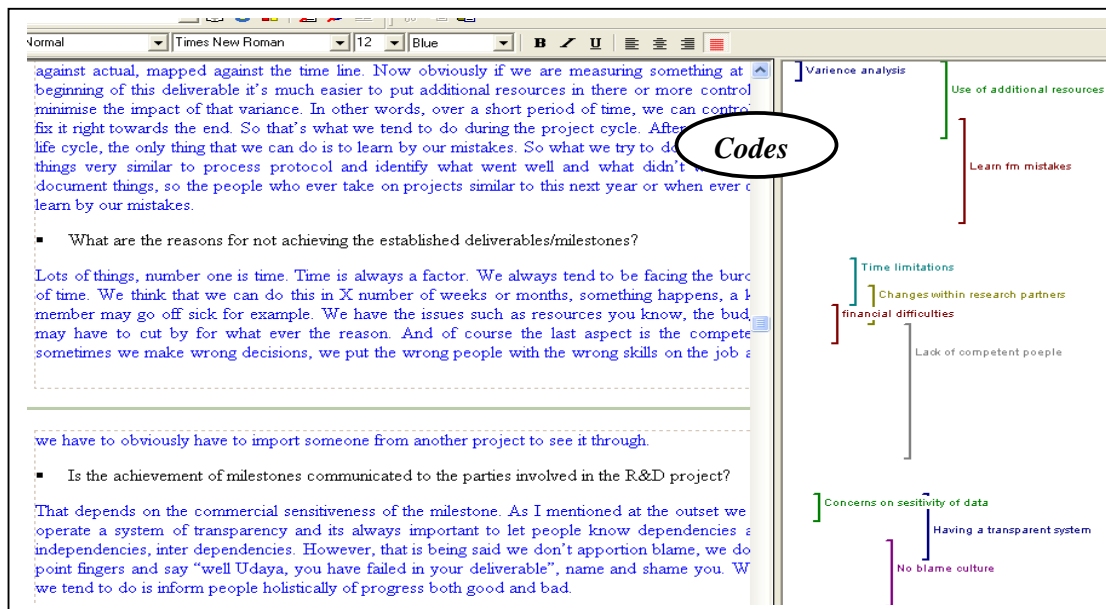

Figure 1: Coding the transcripts

504

The next step was to assign a code for each concept identified from the transcripts. As mentioned in section 3.3, the researchers used the characteristics of both deductive and inductive coding approaches. When a concept was identified, a code was assigned from the provisional list of codes developed through theory or assigned a new code. This process was carried out in repetitive cycles until no new concept is identified. This process enabled the researchers to be more familiar with the interview transcripts and to generate new concepts from the text. The coding was done by using descriptive key words where the key words were chosen from the original words/phrases used in the transcripts. Figure 1 gives an example of a coded interview transcript.

The codes created from the above process were listed as Free Nodes (nodes in a flat structure) in the NVivo software. Thereafter, the next task was to link these codes (which were listed as Free nodes in NVivo) with the research questions to build up arguments and to arrive at conclusions. To do this, NVivo software identifies another type of node which is called as Tree node. Because of the hierarchical nature of Tree nodes, Free nodes can be transferred under them and arrange and structure in an appropriate way. Accordingly, each research question (refer 1.1) was broken down into sub themes and these sub themes were used as the Tree nodes for this study. Having developed the Tree nodes, the Free nodes were matched against them and were transferred accordingly (Figure 2).
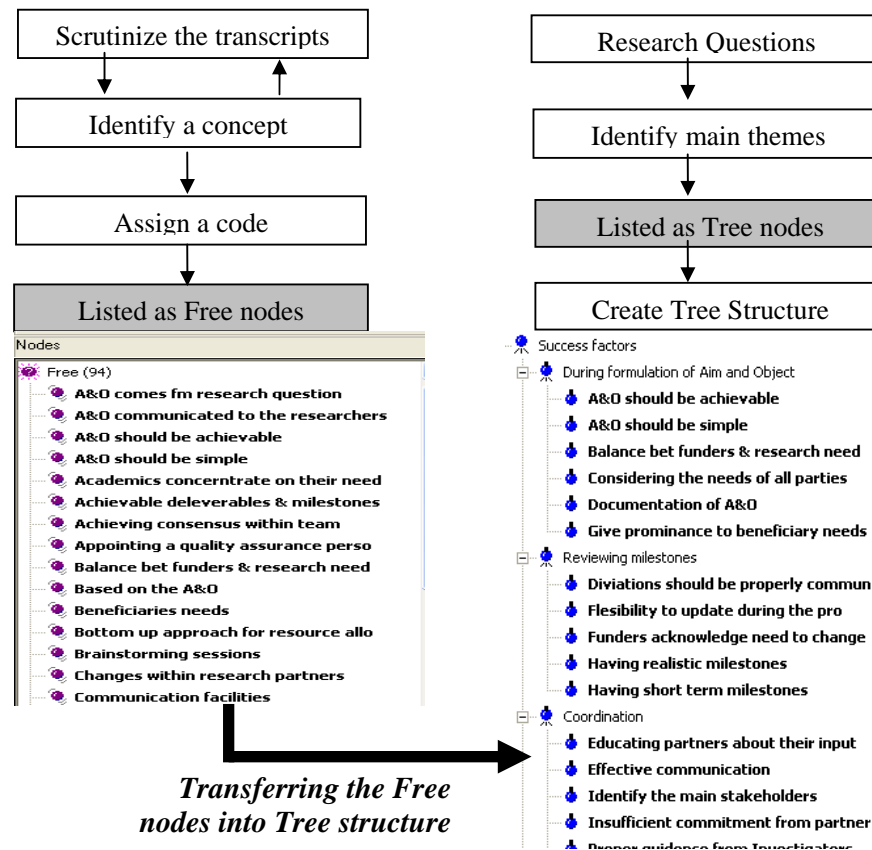

Figure 2: Structuring the codes

Once the coding is completed, it is expected to generate frequency tables for each code. This will be done mainly to fulfill one of the research questions of the study i.e. to evaluate the criticality of the success factors identified from the interviews. In addition to that, the identified codes will be used to prepare a questionnaire survey that will be carried out within

the case study organisations. The un-coded text will be used during the development of arguments linked with the theoretical background derived from the literature support. In addition, the researchers' comments and arguments, together with findings from the literature reviews will be used to further analyse the main concepts and codes derived from the content analysis and finally to fulfill the research questions and thereby to arrive at the conclusions.

5.1 Practical challenges encountered during the data analysis

The comprehensive identification of main concepts is one of the challenges encountered at the data analysis stage. Combination of the deductive and inductive coding approaches, together with repetitive cycles of coding procedure until no new concept is identified, aims to address this issue successfully (see section 3.3).

The words/sentences do not have a single meaning but can be interpreted in different ways when looked at in different perspectives. Thus, the researcher's opinion can influence the codes and categories of the study thereby making them subjective. The researchers therefore had the challenge of minimising the subjectivity of the data set. As Weber (1990, p: 62) noted "Texts do not speak for themselves … the investigator must do the speaking and the language of that speech is the language of theory". Accordingly, to minimize the subjectivity, the researchers pre-established some of the codes using the theoretical background (see section 3.3.1.1). Furthermore, it is expected the research will carry out a questionnaire survey within the case studies and triangulate it with the data obtained form semi structured interviews to identify any discrepancies between the data.

Finally, the researchers encountered the challenge of ensuring the consistency of the codes and categories throughout the data analysis process which is referred to as *reliability* in content analysis. The ambiguity of word meanings or definitions of categories can affect the reliability of the data. During this pilot study, the consistency of coding was maintained by establishing rules to define categories and codes. In addition, the NVivo software assists to increase the reliability of the data through transparency of the analysed data and by avoiding the duplication of codes via its data retrieval facility.

6   CONCLUSION

As explained in section 2, the free flow of ideas through semi structured interviews generated the "unstructured" data; challenging the researchers to structure them. The systematic procedures of content analysis guided the researchers first to *breakdown the textual data into meaningful categories*. Thereafter, the development of codes assists the researchers to *condense the data into analysable units* related to the research questions of the study. Middle range approach between deductive and inductive coding process guided the researchers not only to use pre-established codes to look at specific concepts, but also to generate concepts that were not identified before. Thus, the researchers were able to match theory with the empirical data and had the flexibility of generating new theory that makes the study more complete. Finally, by matching the corresponding codes with research questions the *data was organized* in such a way as to derive conclusions.

The computer software (NVivo) used throughout this process was only utilised to replace the clerical part of data handling whilst the researchers themselves deciding on the concepts, codes and the level of organization of data based on the theatrical background and experience. The use of NVivo software used within this study facilitated rigorous and a transparent study while maintaining its consistency.

However, it should be noted that content analysis has its own disadvantage with regard to its subjectivity and reliability of the data. During this pilot study, the subjectivity of the data was minimized by matching the empirical data with the theory, whilst the consistency of coding was maintained by establishing decision categories and rules to define codes and categories. It can be concluded that use of case study research approach coupled with semi structured interviews facilitated an in depth study to explore the PM concept within the construction R&D activity while content analysis aided the structure of the data in a meaningful way to identify underlying concepts.

7   REFERENCES

Amaratunga, D., Baldry, D., Sarshar, M., and Newton, R. 2002, Quantitative and qualitative research in the built environment: application of "mixed" research approach, *Work Study*, 51.1, pp17- 31

Barrett, S. P and Barrett, L. C. 2003, Research as a kaleidoscope on practice, *Construction Management and Economics*, 21. 7, pp755-766

Berelson, B. 1954, Content Analysis, in G. Lindzey, (Ed.), *Handbook of Social Psychology,* Volume 1, Addison-Wesley, Reading MA

Bernard, H. R. 2000, *Social research methods: Qualitative and quantitative approaches*, SAGE, California and UK

Beuro of Justice Assistance, 2006, content analysis: definitions, (accessed 6th December 2006), available from:
http://www.ojp.usdoj.gov/BJA/evaluation/glossary/glossary_c.htm

Cripps, A., Fidler, J., Lorch, R., and McCaffer, R. 2004, *A research strategy for the construction industry*, CRISP, (accessed 25th December 2005), available from:
http://www.ncrisp.org.uk/Publications/R&I%20workshop%20notes.pdf

CRISP consultancy commission, 1999, *Linking construction research and innovation to research and innovation in other sectors*, CRISP (accessed 15th December 2005), available from: http://ncrisp.steel-sci.org/Publications/ws984_r.pdf

Colorado State University, 2006, *Conducting Content Analysis*. (accessed 6th December 2006), available from http://writing.colostate.edu/guides/research/content/index.cfm

Courtney, R. G. 1999, *Innovative ways of funding construction Research: an ideas paper*, Construction research and innovation strategy panel, (accessed 21st June 2005), available from: http://ncrisp.steel-sci.org/Publications/9913fpRC.pdf

Fairclough, J. 2002, *Rethinking construction innovation, and research: A review of government R and D policies and practices*, Department of Trade and Industry, London

Flanagan, R. and Jewell, C. 2006, Understanding the construction sector - profitability and profit pools, *International conference on building education and research*, 10th -13th April, Hong Kong

Fielding, N. G., and R. M. Lee. 1998, *Computer Analysis and Qualitative Research.*, SAGE, Thousand Oaks, California and UK

Franzosi, R. 2004, Content analysis, in Lewis-Beck, M, Bryman, A and Liao, T F (Ed), *The SAGE Encyclopaedia of Social Science Research Methods*, SAGE, California and UK

General Accounting Office, 1996, *Content Analysis: A Methodology for Structuring and Analyzing Written Material. Report Nr. GAO/PEMD-10.3.1.* United States General Accounting Office, Washington,(accessed 12th January 2007), available from: http://archive.gao.gov/f0102/157490.pdf

Hartley, J. 2004, Case study research, in Catherine Cassell and Gillian Symon (Ed), *Essential guide to qualitative methods in organizational research* (pp.323-333). SAGE, London

Herriott, R. E., and Firestone, W. A. 1983, Multisite qualitative policy research: Optimising description and generalizability, *Educational Researcher*, 12, pp14-19

Holsti, O. R. 1969, *Content analysis for the social sciences and humanities*, Addison-Wesley, MA,

Jorgenson, D. L. 1989, *Participant Observation: A Methodology for Human Studies*, SAGE, California

Kalpan, A. 1943, Content analysis and the theory of signs, *Philosophy of Science*, 10, pp230-247

Krippendorff, K. 2004, *Content Analysis: an introduction to its methodology*. 2nd edition, SAGE, California and London

Kulatunga, U, Amaratunga, D, Haigh, R (2006a), Performance measurement in construction research and development, in Neely A, (Ed) *Performance Measurement and Management - Public and Private*, The University of Cranfield, UK

Kulatunga, U, Amaratunga, D, Haigh, R (2006b), Measuring performance and the impact of research and development in construction: research methodological perspectives, *Proceedings of the 6th International Postgraduate Research Conference*, April 6th-7th, The University of Salford

Mayring, P. 2000, Qualitative Content Analysis, *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research,* 1.2, (accessed 7th December 2006, Available from: http://www.qualitative-research.net/fqs-texte/2-00/2-00mayring-e.htm

Miles, H., and Huberman, M. 1994, *Qualitative Data Analysis: A Sourcebook*, SAGE Beverly Hills

Moscarola, J. 2002, Contribution of qualitative methods to research work and organisational psychology. *Communication to ISSWOV 2002*, Warsaw.

Palmquist, M. 2006, *Content analysis*, (accessed 4th December 2006), available from: http://www.gslis.utexas.edu/palmquis/courses/content.html

Palmquist, M. E., Carley, K.M., and Dale, T.A. 1997, Two applications of automated text analysis: Analyzing literary and non-literary texts. in Roberts, C (Ed.), *Text Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Texts and Transcripts.* Lawrence Erlbaum Associates, Hillsdale, NJ

Patton, E., and Appelbaum, S. H. 2003, The case for case studies in management research, *Management Research News*, *26*.5, pp60-71

Print, M. 1999, *Funding construction industry research and innovation - time for a change,* CRISP, (accessed 25th December 2005), available from: http://ncrisp.steel-sci.org/Publications/9913fpMP.pdf

Robson, C. 2002, *Real world research*, Blackwell Publications, USA

Ryan, G. W., and Bernard, H. R. 2000, Data management and analysis methods, in N. D. Y. Lincoln (Ed), Handbook of Quantitative Research, 2nd edition, pp. 769-802, SAGE, Thousand Oaks

Ryan, G. W., and Bernard, H. R. 2003, Techniques to identify themes, *Field Methods*, 15.1, pp85-109

Silverman, D. 2001, *Interpreting qualitative data: methods for analysing talk, text, and interaction*, 2nd edition, SAGE, London

Stemler, S. 2001, An overview of content analysis, *Practical Assessment. Research & Evaluation*, 7.17

Swan, J. 1997, Using cognitive mapping in management research: Decisions about technical innovations, *British Journal of Management*, 8, pp183-198

Townsend, M. 1999, *Linking construction needs and construction research output: helping supply more effectively meet demands*, CRISP, UK

Weber, R. P. 1990, *Basic content analysis*, 2nd edition, SAGE, London

Wikipedia, 2006, *Content analysis*, (accessed 7th December 2006), available from:
http://en.wikipedia.org/wiki/Content_analysis

Yin, K. 2003, *Case study research: Design and methods*, 3rd edition, SAGE, London

Zonabend, F. 1992, The monograph in European ethnology, *Current Sociology*, 40.1, pp49-55

Zhang, T., and Kuo, J. 2001, Audio Content Analysis for Online Audiovisual Data Segmentation and Classification, *IEEE Transactions on Speech and Audio Processing*, 9.4