On the use of Hidden Information as a Measure of Complexity in Supply Chains.

Peter Dickinson
00063665
Salford Business School, University of Salford, Manchester, UK.

Submitted in Partial Fulfilment of the Requirements of the Degree of Doctor of Philosophy, June 2015.

Table of Contents

Tables

Table of Figures

Table of Tables

Equations

On the use of Hidden Information

Abstract

With the increase in globalisation, mass customisation becoming ever more ubiquitous, product life cycles becoming shorter, midlife upgrades becoming more popular; with some product based solutions transitioning to service based solution, the supporting industry supply chains are becoming ever more complex. There is a likelihood that this complexity will increase with increasing access to emerging market cost effective capabilities and an international customer base. Working with this complexity is one of the challenges facing the supply chain executive, and there is no reason to think this increasing complexity will go away over the coming few years. Creating and managing an effective supply chain structure will be a key performance target and potentially a key differentiator. One element of the management task will be the understanding of structure and how it impacts on the overall key performance indicators.

Structures can be represented as activities and connections. These structures can be set out to reflect the complexity of the structure necessary to cover all potential business scenarios. A key determinant of a structure will be how ‘mixed up’ it needs to be to cope with the demands of complexity, flexibility and agility necessary for all the business scenarios. It is possible these structures can be represented as a matrix and, using information theory, analysed to measure complexity. This thesis looks to use a matrix approach and address these challenges by offering a revised model for structural complexity in the supply chain. Like most research in this field, this thesis will be experimental and laboratory based; however, the scenarios used in the analysis will be validated externally.

The aim of this research is to make a contribution to the research in this field by distinguishing between complexity, variability and structural complexity; providing a

framework and quantifiable measure of complexity for a supply chain governance structure using information theory and graph theory; analysing the impact of language aggregation on the hierarchical business process. Additionally the research assumptions have the aim of making this research practical for the management practitioner.

1 Chapter One: Introduction

Supply chain management, in some industries, has struggled to gain recognition as an important aspect of business operations. In industries such as food and automotive for instance, the concept of a supply chain is embraced and recognised as a key contributor; whereas in other industries, military aircraft or heavy industry perhaps serve as examples, the concept wrestles with traditional methodologies and transactional interactions that are classified, conceptually, as supply chain management.

The current transitions in the economic climate: rising debt levels and spending reductions across the western hemisphere and growth in the Indian, Chinese and Brazilian economies, for instance, add to the globalisation and/or internationalisation pressures on organisations' supply chains. Internal to the organisation this pressure is only exacerbated: increasing levels of customer sophistication, mass customisation, the ever decreasing life cycle of most products all add to the challenge faced by supply chain management. The combined internal, external, product and market challenges only add to the complexity of form in the supply chain necessary to rise to these various challenges.

Complexity of form in the supply chain is simply - and superficially - evidenced through the increasing complexity involved in the definition of the metaphor. This will be discussed in more detail later; however, for now, this increase in complexity still suffers from a general tendency to take a reductionist scientific perspective on analysis. As Gharajedaghi (2011) explains in his excellent book, Systems Thinking: Managing Chaos and Complexity: love does not exist because of the existence of a heart, lungs, liver, brain etc. It does not have a colour, smell or sound; but it does exist as a second order (he uses the term emergent) property of the system. The increased interest in a systemic view; looking from the environment, system and sub-system perspective, adds a new dimension to the challenge of understanding complexity of form. While the metaphorical heart, lungs, liver and brain of a supply chain have been well researched, research on the supply chain as a system; a system of internal, external, product and market interactions, is less pronounced and there is a need to add to the existing research on supply chains as complex systems.

Complexity theories have been researched and applied to various problems, complex adaptive systems for instance, and it has been this general acceptance of non-linearity and human sense and response that have highlighted the importance of research concerning the application of complexity theories to organisation problems. This human element of complexity brings with it the challenge of understanding what the complex problem is; the sense that complexity is about trying to understand a given problem or situation is one of the focus elements of this research. Complexity of form in the supply chain could then be concerned with how the form of the supply chain is understood: What does it take to understand the form of a given supply chain? Recognising the systemic perspective, this thesis will develop an approach to understanding complexity in the supply chain.

To do so, this thesis will review the existing research on supply chain related complexity. Using primary theories from this research; specifically, entropy, information theory and network theory the thesis will synthesise the research and propose a revised approach and method for understanding the complexity of form of the supply chain. In doing so, the intention of the research is to develop an approach that is simple enough to be useable in an operational context.

1.1 Background

Entropy, in the physical sciences, has generally been considered to be the energy not available for work, expressed as a measure of disorder. In statistical mechanics, entropy is considered to be the amount of additional information required to understand the state of the system. In Information theory, entropy can be defined as a measure of the uncertainty associated with a set of data. The usefulness of entropy as a measure of uncertainty in various aspects of business has received some attention so far in the academic literature (Ebrahimi, Maasoumi, & Soofi, 1999; Blecker, Kersten, & Meyer, 2005; Calinescu, Efstathiou, Sivadasan, Schirn & Huatuco, 2000; Cardoso, 2005; Dionisio, Menezes, & Mendes, 2005; Efstathiou et al., 1999; Frizelle, 1998; Frizelle & Woodcock, 1995; Gleik, 2011; Gonzales, Rubio, Gonzales, & Velthuis, 2010; Jung, Chin, & Cardoso, 2011; Karberger & Masoon, 2001; Kumar, 1987; Li & Vitanyi, 2008; Rao & Gu, 1994; Lassan & van der Aalst, 2009; Scuricini, 1988; Shannon, 1948; Shuiabi, Thomas, & Biuyan, 2005; Sivadasan, Efstathiou, Frizelle, Shirazi, & Calinescu, 2002; Wilding, 1988; Yao, 1985). These researchers have looked, collectively, at the following important issues:

- Differentiating between the structural and dynamic types of complexity in the business.
- Justifying the information theory variant of entropy as a valid measure of supply chain complexity.
- The creation of frameworks and methods for measuring complexity in the supply chain.

On the use of Hidden Information

All of the referenced research involving entropy as a measure of complexity use Shannon’s (1948) information theory variant as a definition. The validity of which will be considered in more detail later.

In the physical sciences entropy has developed in two forms: macro level entropy developed by Clausius (c.1850) and micro level entropy developed by Gibbs and Boltzmann (c.1867). The relationship between the micro level physical science variant and the information theory variant can be traced back to Shannon’s 1948 research. To be specific, although almost identical in form, the Gibbs – Boltzmann and Shannon equations were developed separately. Folklore has it that it was only after Shannon developed his research – aparently during a conversation between Shannon and Jon von Neumann – that the similarity in forms led to the label ‘entropy’, already applied to the Gibbs – Boltzmann variant, being applied to the Shannon information theory variant. The validity of this labeling and the form of the formula have already been questioned to some extent (Ben-Naim, 2011; Tsallis, 2009). Given this, the assumption that Shannon’s variant of entropy as a measure of uncertainty in the above research should be revisted; consequently, this research will review the Shannon variant for entropy and propose a modified approach relevant to the understanding of complexity in the supply chain.

Research on graph theory has developed with the increase in connected networks and the use of computers. Parallel to this, some industries have recognized the integrated set of organizations that convert and add value to raw materials; i.e. the supply chain, can also be viewed as a network of interactions. Thus the representation of a supply chain as a network of interactions can be studied from the point of graph theory. In fact, this adds another dimension to the current thinking of supply chain management as a set of interacting processes in that with graph theory, linkages as well as activities are considered. Furthermore, the study of supply chain management as a network is essential: Dorogovtsev (2010) argues that specific network architectures, namely cage graph architectures, in synchronised systems - like a supply chain - offer an optimal architectural design. That said, graph theory, in the context of understanding supply chains does have limitations; these will be explained later in the thesis along with an alternative approach to the use of graph theory in understanding supply chain complexity.

The origins of information theory appear to be attributed to Ralph Vinton Lyon Hartley of Bell Laboratories who, in 1927, published what appears to be the first seminal work on the subject. This was followed, in 1948, by the work of Claude Shannon, also from Bell Laboratories, who published what has become the de facto standard for formulating the understanding of complexity in information theory: Entropy. It is worth noting that Hartley did not use the term entropy, and his formulations for information theory did not align with the physical sciences or Shannon formulations. Shannon’s 1948 work on information theory is not without its critics: In 2009 Tsallis questioned the form and universality of the phenomena labeled ‘entropy’; offering a non-extensive version as an alternative. In 2011 Ben-Naim questioned the formulation and labeling for what Shannon termed entropy; proposing alternative formulations and a different name for what entropy is supposed to represent. Thus it makes sense in this research not to accept – as has been the case with other research in the field – Shannon’s version as the de-facto standard.

The intention of this thesis therefore is to synthesize the works on entropy, graph theory and information theory to propose an approach for understanding complexity of form in supply chains. In doing so the purpose of this thesis is twofold: Firstly, to add to the existing research on the application of entropy as a measure of uncertainty in the supply chain, and more specifically, to explore the relationship between entropy as a measure, and the organisational structure of the supply chain. Secondly, to explore the use of an alternative method for capturing uncertainty and complexity in the supply chain. In doing so the research will not lose sight of the issue of actor legitimacy as an input to the sustainment, addition to, or removal of uncertainty and complexity in the supply chain. The hope is that this research will contribute considerably towards advancing the current state of knowledge in this area, and provide an approach that is easy to use from a practitioner perspective.

1.2 Motivation

Supply chain management has been variously defined, and well researched, as a set of individual interactions that cause the movement of material, information and cash between the contributing organisations. The study of organisations as a system has also been well researched (Ackoff & Emery, 1972; Sterman, 2000; Beer, 1994; Gharajedaghi, 2011); however, with a few exceptions (Sterman, 2000 and Streetfield, 2001, for instance), specific research on supply chain management as a system is less well researched. The concept of supply chain management as a system is a broad subject when considered from the four facets, suggested by Gharajedaghi (2011), to be the core components of system thinking. It is not intended herein to cover each of these components; however, the motivation for conducting the research is to contribute to the systems thinking aspect of supply chain management.

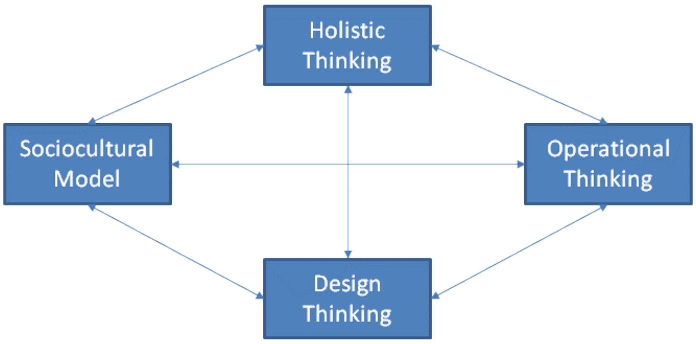


Figure 1: Foundation of Systems Thinking (adapted from Gharajedaghi, 2011)

The motivation for the thesis – thinking about supply chains as a system – calls for a research strategy and design that recognises the systemic nature of the topic. Because of this systemic view, the next few section will take some time to outline the issue of research strategies before defining the specific strategy and design to be used.

1.3 Relevance

The research is relevant, primarily, for five reasons: It questions the generally accepted use and validity of entropy as a measure or proxy for complexity. In doing so it offers an alternative measure for complexity in organisation structures. The research is further relevant because the solution offered is reasonably practical; determining an alternative measure by analysing the governance structure of the organisation. And when analysing the governance structure, the research is particularly relevant because it conceptualises and measures in two dimensions. Finally the research is relevant because it conceptualises the process of research as being similar to the two dimensional content of the research in the construct of ontologies.

1.4 Aim and Objectives

The aim of this research is to synthesis the use of entropy, network theory and information theory as an approach for understanding uncertainty - as complexity of form in structural components in a supply chain system.

The objectives of this research are five-fold and can be summarised in the following points:

- To critically review the existing literature to appreciate the various approaches in this domain.
- To advance the current state of understanding in the area by introducing an amended approach to conceptualising complexity in the supply chain.
- To investigate and compare the use of entropy as a measure of complexity, with a revised measure proposed herein.
- To reflect on the experimental findings and comment on the role of the actor in upholding the structural perspectives.
- To develop an approach that is reasonably practical to implement.

1.5 Research Questions

Accepting the general increase in complexity in business and supply chains specifically, the research question considered herein is: Can a measure of hidden information be used to quantify structural complexity in supply chains? Chapter two will explain the background and build the process for creating hidden information as a measure and chapter three will provide a fuller description of the research question.

1.6 Contribution

The contribution to theory and practice made by this research can be summarised in five distinct points. The existing literature broadly relates complexity with variability; that is, variability is a component of the complex environment. This research looks to separate out variability from complexity. The reasoning for this is that variability is an everyday occurrence in business; the management of which is a pre-requisite of business operation. Variability has an impact on the parameters of process execution and, as a consequence, the operational complexity. It does not have an impact on structural complexity unless the variability exceeds the structural components of the supply chain; in which case the structure and operations become chaotic; a higher order complexity. A framework that enables the understanding of the extent of supply chain process and parameter dispersion necessary– the variability - to govern and operate a supply chain is required.

Research into the use of information theory and entropy as a measure of uncertainty has been carried out in some operational domains, and there appears to be little research on the application of the two theories in a supply chain context. The purpose of this research is to further contribute to this field of research.

Supply chains are under increasing pressure to operate in complex environments. It is perceived that adding to the body of knowledge in this area could be timely in terms of business interest and application. Furthermore, the development of an approach to the measurement of the dispersion of variability as an extension to an already popular model may improve the data feedbacks and application.

On the use of Hidden Information

The business process model, from a non-information technology perspective, does not bring with it the logic gate construct necessary for an IT workflow construction. Such IT workflows offer a measure of complexity as a function of the logic gate construction. Where the business process model is not part of a logic gate construct, but is part of a business governance model, the ability to quantify a method for understanding complexity in business processes will contribute to the effective application of business processes to business operations.

From a research perspective Hartley (1927) explained the relationship between the length of the message, the language and the amount of information communicated; that is, as the number of language symbols reduce, for a give piece of information, the length of the message increases. Recognising this, there is a need to understand the effect on information brought about by the aggregation of the original language into a summarised language. It follows that if a language is aggregated into a lesser language, more symbols would be required to describe the state of the systems, and yet, in the case of management information, the structured approach assumes information can be aggregated into a lesser language without the loss of information. This seems contrary to Hartley (1927); hence there is a need to understand the effect on information, brought about by the aggregation of a language into management information. Understanding this contradiction will be a further contribution.

Finally, apocryphal, anecdotal or otherwise, the story surrounding the label ‘entropy’ defined by Shannon and associated with the form $-p \log p$ for information theory has been questioned. This research contributes by further analysing and questioning this association.

1.7 Method

The research will take a generally positivistic approach to analysing the process structures of organisations using a common process language (SCOR^{TM}) and using entropy as a measure of complexity. Initially, and for the purposes of this thesis, the research will be experimental; however, it is hoped that, post this piece of work, the approach can be developed and tested further through data provided either directly or via online data entry that will be developed and made available at www.supplychaincomplexity.co.uk. The latter approach is the outcome of the researcher's efforts to establish a dedicated site to facilitate further research in this area. Throughout the research a thread of critical realism is maintained, sometimes the texts winds and twists in support of this epistemology; for instance, chapter four takes some time to contextualise and define fully the research ontology and epistemology. The author askes for your indulgence in this section, it does eventually contribute directly to the research findings.

1.8 Thesis Structure

The remainder of this research is set out as follows: Chapter 2 sets out the theoretical foundations for the research; focussing on entropy, information theory and network theory, before developing a revised model. Chapter 3 is a short section that sets out the research question; which is followed by Chapter 4 that takes time to set out the issues of research and the methodology. The analysis and discussion are presented in Chapters 5 and 6 respectively. Chapter 7 draws conclusions from the research.

As was suggested above, the structure of the research, at times, takes time to set context before moving on; hence, from time to time, the text winds and twists through four topics – complexity, entropy, information theory and network theory. Figure 2 provides a view of this journey, the reader may want to come back to this diagram from time to time.

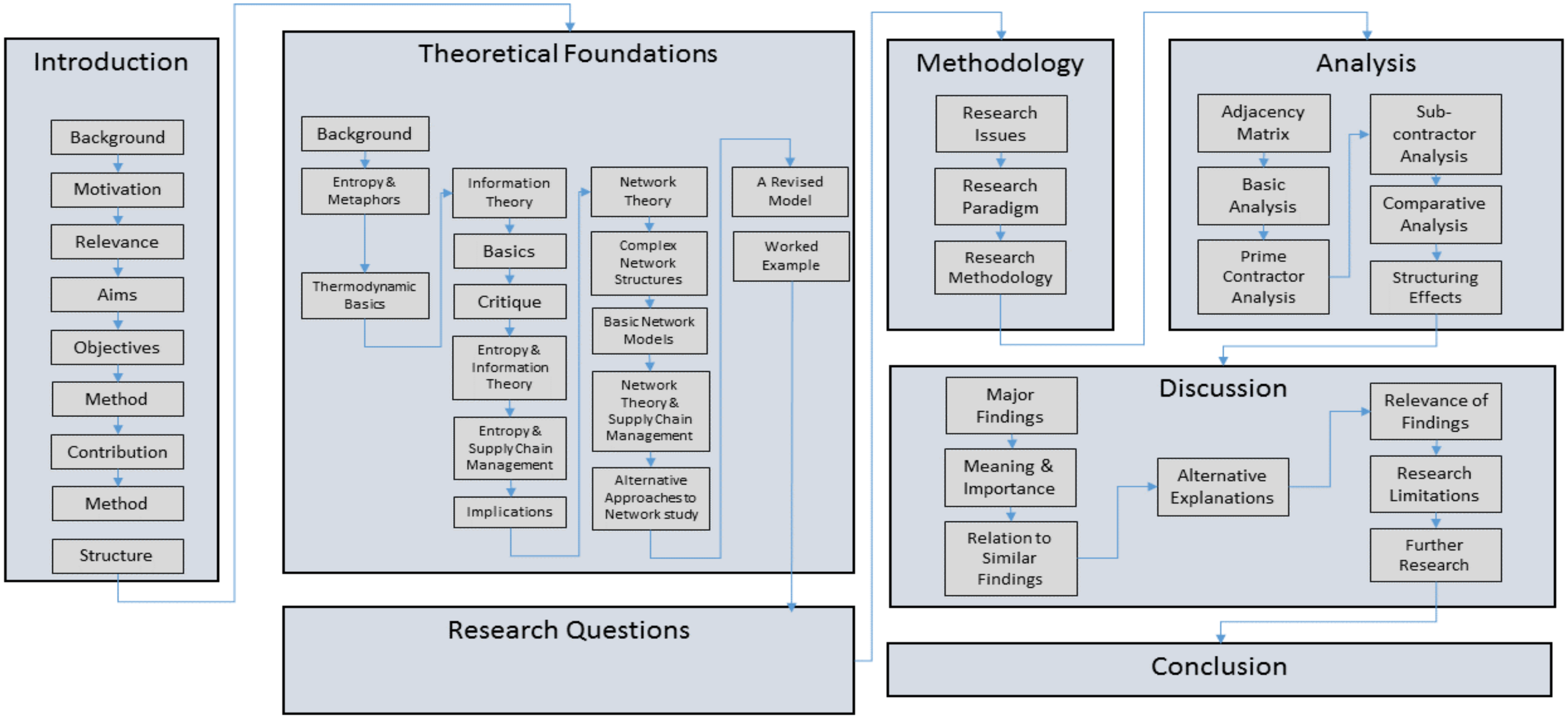


Figure 2: Structure of this thesis

2 Chapter Two: Theoretical Foundation

2.1 Background

Supply Chain Management, as a concept, has been variously described and defined. Giannakis, Croom, and Slack (2004) provided the following sample of definitions. Oliver and Weber (1982), Supply chain management covers the flow of goods from supplier through manufacturing and distribution chains to end user. Jones and Riley (1987), Supply chain management techniques deal with the planning and control of total materials flow from suppliers through end users. Ellram (1991), An integrative approach to dealing with the planning and control of the materials flow from supplier to end user. Harland (1994), Supply chain management is defined as the management of the flow of goods and services to the end customer to satisfy their requirements. Berry (1994), Supply chain management aims at building trust, exchanging information on market needs, developing new product, and reducing the supplier base to a particular original equipment manufacturer so as to release management resources for developing meaningful long term relationships. Cooper et al. (1997), An integrating philosophy to manage the total flow of a distribution channel from supplier to ultimate customer. Lee and Ng (1997), the management of a network of entities that start with the suppliers' supplier and end with the customers' customer for the production and delivery of goods and services. Handfield and Nichols (1999), the supply chain encompasses all activities associated with flow and transformation of goods from the raw material stage, through to the end user, as well as associated information flows. Material and information flow both up and down the supply chain. Supply chain management is the integration of these activities to achieve sustainable competitive advantage. Simchi-Levi (2000), Supply chain management is a set of approaches utilised to efficiently integrate suppliers, manufacturers, warehouse and stores, so that merchandise is produced and distributed at the right quantities, to the right location and at the right time, in order to minimize system wide costs while satisfying service level requirements. Ayers (2001), Supply chain management is the design, maintenance and operation of the supply chain processes for the satisfaction of the end user.

The various definitions of supply chain management have continued to mature: Emmett and Crocker, (2006) define supply chain management as ‘The process that integrates, coordinates and controls the movement of goods, materials and information from the supplier through a series of intermediate customers to the end customer’. In some cases the definition is contingent on other business activities. Wang et al. (2007), using the work of Narasimhan and Mahapatra (2004), take this perspective by suggesting 'a supply chain is established when there is integration of operations across its constituent entities, namely, the suppliers, partners, and business customers.' Another alternative perspective is that of defining what a supply chain is rather than what it does or what is the situation in which a supply chain emerges. Bowersox, Closs and Cooper (2007) took this approach by suggesting supply chain management is an amalgam of organisations collaborating to leverage their position and improve operating efficiency. For each firm involved, the supply chain relationship reflects a strategic choice.

.

The definitions and descriptions of supply chain management have also considered scenarios wider than the ‘functional’ perspective of what a supply chain is or does. These wider scenarios consider the supply chain to be part of a network; for instance, Harrison and van Hoek (2008) state: ‘A supply chain is a network of partners who collectively convert a basic commodity into a finished product that is valued by the end customer, and who manage returns at each stage’. Continuing with the recognition of supply chain management as more than a functional construct Sun and Ye (2008) assert the flow of information, logistics and funds flow are the ‘surface phenomena’ of the supply chain, the outcome of a more fundamental set of underlying elements of information, human resources and organisational structure. The impact of organisational structure of supply chains will be a feature of this research.

Supply chains as a structure are further complicated with the recognition that they act as a consequence of behavioural free will; the structure needs human intervention to make it work. This perspective has been noted relatively recently by Sokolov and Ivanov (2010) who take a complex adaptive systems perspective by recognising the complex dynamics and behavioural influence: 'Supply chains may be justifiably called complex dynamic multi-structural systems with active elements of free will behaviour'. Whilst this perspective is only recent to supply chain management research, it has been previously researched in the wider management domain; for instance, Dooley (1997) defined complex businesses as 'semi-autonomous organisational members interacting at many levels of cognition and action.' Given these varied and maturing definitions of supply chain management, for research in the subject to be succinct, a specific definition for the supply chain management metaphor relative to this research is required. This will be developed later in the text.

The development of the perspective of supply chain management as a complex systems has continued to mature, sometimes as supply chain management research and sometimes as complex adaptive systems or systems dynamics research, using supply chain management as an example. Erenguc, Simpson, and Vakharia (1999) assert supply chains can be defined using two components: (a) the supply chain network and, (b) the nature of the relationships between each stage of the network. The behavioural aspect of this network relationship has been highlighted by, for instance, Ivanov and Sokolov (2010) who present the concept of agent behaviour in the supply chain. Taken from the research on complex adaptive systems, their approach recognises the behavioural impact of the agent on the network. The concept is not new. Sterman (2000) in his book Systems Dynamics uses a supply chain as an example of the complex systems interplay of a business. While complex, the advantages of such an approach were highlighted by Cousins and Menguc (2005), who extended the work of Barney (1991) but contradict the work of Dierickx and Cool (1989) in asserting that strategic advantage in an inter-organisational context can be accumulated outside the individual organisation. The research on organisations as complex systems continues, see for instance, Stacey (2010), Shaw (2001) and Streatfield (2002) are interesting examples, but their specific areas of research are beyond the scope of this thesis.

The work on transaction cost economics is particularly pertinent to the wider supply chain conceptual structure. Williamson (1975) first explained the transaction cost concept in terms of four factors: bounded rationality, asset specificity, the concept of opportunism and imperfect imitability; centred on a structure of contractual obligations and vertical integration. In 2002 he revisited the transaction cost concept and proposed the principle of forebearance whereby there is an acceptance that contractual obligations are not fixed and that there is a different principle of obligation. These two views represent the inter-organisation relationship dimension in the conceptual understanding of supply chain management, with arm’s length and collaborative relationships positioned at either end of the dyadic. Day and Nedungadi (1994) support the suggestion that supply chain management is as much about relationship and, at a deeper level, a shared mental model of the shared knowledge of order, logic and relationships. Welch and Wilkinson (2002) take this suggestion even further by asserting that such mental models or schemas are essential determinants of the supply chain construct, enabling managers to make sense of the interaction. In this case the schema is becoming the cognitive map which provides each individuals’ boundedly rational backdrop against which the individual and group interplay dynamics operate.

Thus far dimensions of material, information, data, structures and behaviours have been described as being key to a succinct definition of supply chain management. Before a research specific definition can be attempted, there appear to be three further aspects to consider. Firstly. Dooley (1997), referring to the work of Ackoff and Emery (1972) and Hayles (1991), asserts the management paradigm of any given era can be aligned with the prevailing era's scientific lead paradigm; a model that suggests, for instance, that structural aspects of the organisation are contingent on research on organisation structures. Similarly, Emmett and Crocker (2006) relate changes in the supply chain management paradigm to the parallel changes in technology and business approaches. It follows that cognisance of current research trends and topics will add a further dimension to any definition. Secondly, an organisation ecology view - where the attributes of the organisation are tested in a Darwinian paradigm, with the 'fittest' surviving - has been suggested by Hannan (Hannan & Freeman, 2004). Importantly, this view captures the life cycle of the organisation and attributes therein, i.e. the simple need for ongoing innovation in, and transformation of, a business in order to sustain a competitive market position. Any definition should recognise this need for continuous evolution of an organisation's attributes.

All the above, rather than being seen as a list, represents a journey from the early basic definitions of supply chain management through to a point where supply chain management is increasingly recognised as a complex multidimensional concept does not align with the traditional view of business functions; hence the third and last dimension to add to this increasingly complex definition is the cognitive view. This view, mentioned in the discussion on transaction cost economics, develops a position broadly aligned with the Frankfurt school and specifically the work of Lukas who argued the difference between the morally neutral scientific technocratic rational institution and social contingency (Feenberg, 2014), or put simply, individuals are only boundedly rational and only able to view things from their own cognitive perspective.

From all the above a revised definition of supply chain management is offered:
Based on the individuals' perceptions, research and understandings, the complex interplay between inter and intra organisation structures that systemically move material, information, data and funds between entities may be called a supply chain.

Interestingly, a level of complexity, interpreted by actors using only boundedly rational logic, is not a new issue. As far back as 1974 Galbraith discussed the manner and methods adopted by organisations coping with complex or uncertain environments. Specifically, reducing the need for information, increasing the capacity for information acquisition, storage and retrieval. There is a need to be certain on the concepts raised by Galbraith; complexity is a wide and variously studied subject; equally, risk and uncertainty are intertwined as concepts that have different meanings. Coase (1937) referring to Knight (1933), suggested risk to be something that a decision maker can assign a mathematical probability to; whereas uncertainty cannot be expressed in such terms. Keynes (1937) reiterated the same sentiment; hence a measure of complexity related to the measureable components necessary to manage uncertainty would be useful when considering supply chain structures.

2.2 Complexity

Complexity, as a metaphor, brings with it an array of definitions and interpretable spaces (Gonzales, Rubio and Gonzalez, 2010). From conceptual discussions on the metaphor (Ashby, 1956), through many domains, to specifics such as, for example, discussions on complex adaptive systems (Bennet and Bennet 2004), information theory (Reza, 1994), agent modelling and behaviour (North & Macal, 2007), and business dynamics (Sterman 2000). Stacey (2010) asserts there to be no single science of complexity, but all complexity sciences contain the problem of non-linear interactions which, in turn, leads to the problem that the models for such systems cannot be solved. The removal of the direct cause and effect relationship resonates in organisations where human interaction and interpretation create a nonlinear element to the organisations capability (Prigogine 1997.) From a supply chain perspective, complexity has been considered from a number of views.

Scuricini, (1988) stated: ‘Complexity is a subjective quality, its meaning and its value change following the scope of the system being taken into consideration’. Milgate (2001), in the context of supply chain complexity, argued that complexity should be viewed as the deterministic component more related to the numerousness and variety of the system. Blecker et al. (2005) took a disciplinal approach stating that complexity research can be found in systems theory, cybernetics, chaos theory and information theory. Wilding (1998), explains business complexity as a triangular concept consisting of deterministic chaos, amplification and parallel interaction. Milgate (2001) explains supply chain complexity as a combination of uncertainty, technological intricacy and organisational systems. None of which should come as a surprise given Williamson's (1975) transaction cost economics view on asymmetric advantage; business ultimately, is simply the exchange of value across an asymmetric state in a complex system. From the above there appears to be two distinct views on the complexity metaphor. One view considers the term to refer to a subjective sense; the other a deterministic sense. This is not unlike the Keynes type model of things that are not known - the subjective component, and things that are known but variable – the deterministic component. The subjective component, while valid as a strategic risk component, cannot be considered as an operational component of complexity; thus, the definition of complexity here will refer to the metaphor in the input deterministic sense; that is, this is a complex non-linear system. The inputs may be known, but the outputs are indeterminate and probable.

A further field of research has been the application of complex adaptive systems theory to the supply chain domain. Choi, Dooley and Rungtusanatham (2001) conceptualised supply networks as complex adaptive systems, proposing ten propositions for understanding the complex perspective. From the supply chain management perspective Sivadasan et al. (2002) suggest the interacting network of a supply chain can be inherently complex. Their definition of complexity synthesises the work of Frizelle's (1998) two class complexity construct with the entropic view of Calinescu et al., (2000), which states that a systems entropy represents the amount of information required to describe the state of the system. The combination of the two perspectives suggests that operational complexity can be quantified as the amount of information required to monitor the state of the system.

Using the work of Frizelle and Woodcock (1995), Sivadasan, Efstathiou, Calinescu, and Huatuco, (2006), distinguished between structural and operational complexity. Choosing operational complexity as a focus for their research, they assert that complexity can be associated with uncertainty of information and material flow within and across organisations. They go on to suggest that the operational complexity of the system, and hence the amount of information required to describe the state of the system, can vary with volatility of customer demand, reliability of supply and internal performance. This assertion appears only to be partially correct: Complexity, in a general deterministic sense could refer to either the structural and/or the operational element of the supply chain. Volatility of demand, supply and performance is an issue of variance rather that complexity; for instance, inventory safety stock is a function of demand and supply variance more than the complexity of the operation; equally, re-order points will be a function of complicatedness in the organisation and process structure. In this research we intend to recognise operational complexity as having a given amount of variation necessary for the smooth operation of the business; therefore operational complexity is defined as the dispersion of information and data across the structure of the business. Frizelle and Woodcock (1995) argued that structure comes before operations, for an organisation to remain in control. This is obviously correct. The organisation governance structure, functions, processes and parameters need to be in place before an operation embarks on throughput; otherwise the operation is simply not in control. If operation were to reach this point they would become analogous to the subjective component described above. Complexity then refers to a determinist component found within the structure of the supply chain, the complexity of the functions, processes and parameters necessary to keep the operational element in control.

For a system to remain in control, and not chaotic or stopped, the necessary structure in terms of processes and governance rules for a given set of data parameters must be in place prior to the operational system being enacted. Re-stating Frizelle and Woodcock (1995), the structural complexity needs to be established before operational complexity occurs. The structural complexity of the system should encompass two elements, that which covers the systemic governance processes and rules and that which specifies the range of allowable values for each of the data elements specified in the processes and rules. This research takes this systemic perspective defining structure as the governance processes, rules and data parameters necessary for the controlled operation of the business. Diagrammatically, the rule, process and parameter, structural, and operational complexity can be represented as multiple interacting structural groupings that bound the governance of the business.

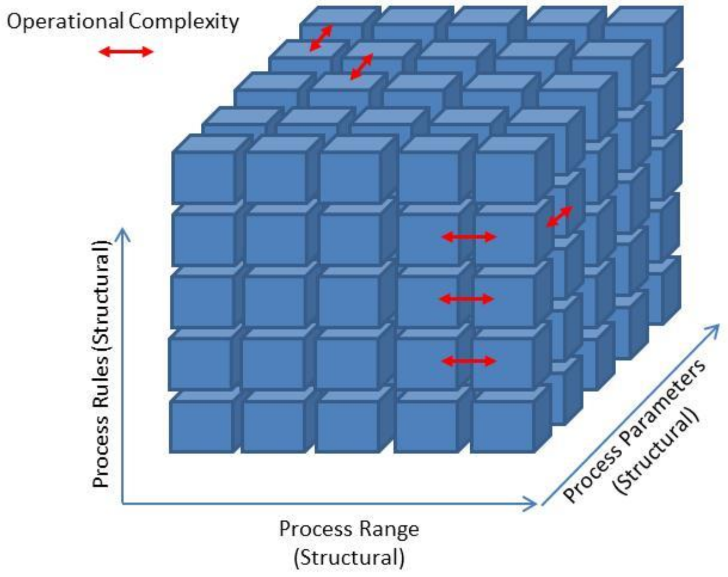


Figure 3: Structural and Operational Complexity

Information theory argues that as a system becomes more uncertain, it becomes more complex and, therefore, more information is required to describe the state of the system. Blecker et al., (2005) assert the increase in complexity on the supply chain as being driven by mass customisation that increases the diversity with which the supply chain needs to cope, and hence creates an increase in the level of structural, process and data complexity to support such diversification. Thus, complementary to Sivadasan et al., (2006) a measure of complexity derived from the amount of information required to describe the state of the supply chain and the degree to which the information and data is dispersed throughout the supply chain structure, such that the system is kept stable; i.e. the supply chain continues to perform operationally, would be valuable. The advantage of this approach is that it is not subject to the variations in the data set for any given time period as would be the case in the approach taken by Sivadasan et al. The approach defined herein therefore aligns with a definition of complexity as being a measure of the dispersed information required to describe the state of the system as a function of the business process model required to cover all business scenarios within the boundary of each business capability, and the range of data required for such management. Complexity in this research is therefore considered to be a measurable dispersion of the necessary data required to manage and control the information that ensures the operational flow through a multi enterprise supply chain solution.

Understanding process measurement is important to an organisation simply because processes are the generator of costs and revenue via the discharging of tasks to produce products or services (Jung et al., 2010; Gonzalez et al. 2010). Jung et al. (2010) defined business processes as a collection of tasks and decisions to produce products or services in an organisation. The measures of which are defined as complexity (Cardoza, 2005; Rassen and van der Aalst, 2009) and density (Medling; 2006; Reijers & Venderfeesten; 2004). An organisation then can be defined as a structured set of processes and parameters through which information and data are distributed. If the set of processes x^1...x^n are the processes necessary to govern the tasks of the business and the parameters y^1 ... y^m are the set of parameters available to the x^1...x^n processes, the structure x^1y^1 ... x^ny^m could be laid out as a matrix A. Simplistically this could be represented diagrammatically; an example is laid out in Figure 4.

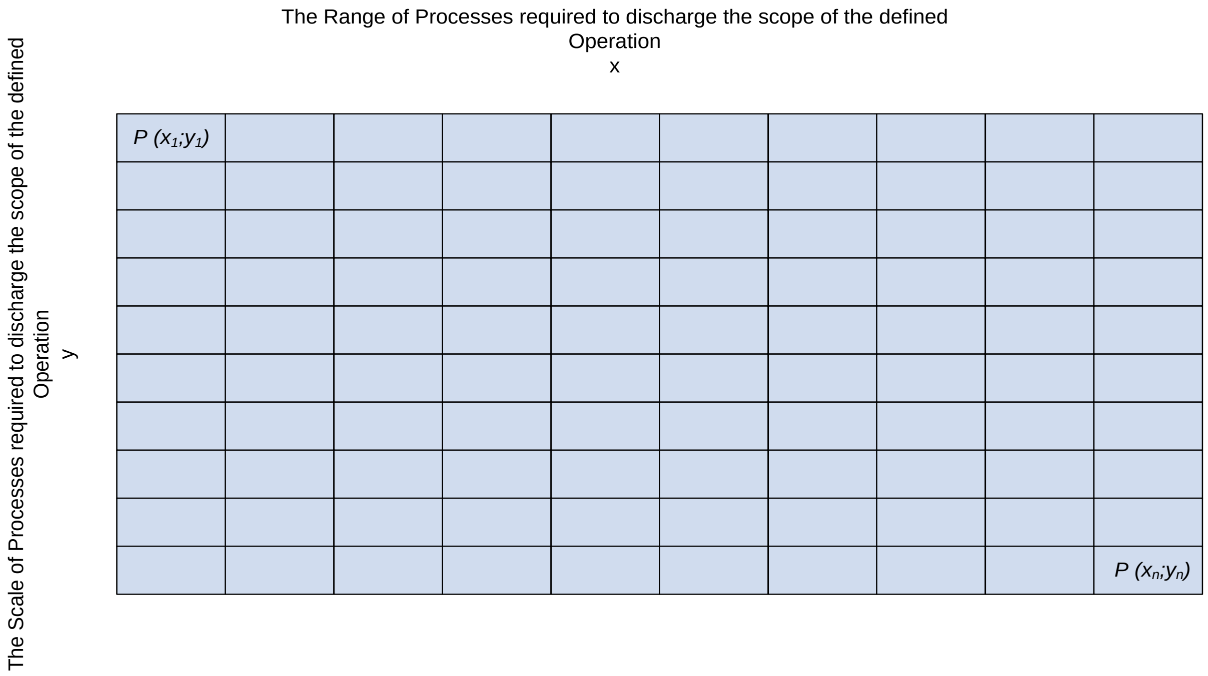


Figure 4: Diagrammatic representation of the range and scales of supply chain processes necessary to operate the business.

Where $x_n \in X$ are the processes from the set of processes X necessary for the controlled operation of the defined business, and $y_n \in Y$ are the parameters from the set of parameters Y necessary to limit the range of allowable values in the processes x_n such that all the allowable business scenarios and data elements can be managed. As such, the structure of the business equates to the range and scale of the defined processes:

$$\sum (x_1 y_1 \ldots x_n y_n) \qquad \text{Equation 1}$$

Not all of the processes defined at any intersect $x_n y_n$ will be used with equal probability; that is, some business processes will be utilised more than others in the discharge of the business activity. Let P_k be the probability that processes $(X\colon Y) = \sum(x_1 y_1 \ldots x_n y_m)$ will be utilised in completion of a business activity. The process structure of the business can therefore be mapped against a matrix, similar to Figure 4 and inclusive of volume and derived probability such that.

$$P(X;Y) = \sum(p[x1;y1] \ldots . p[xn;yn]) = 1 \qquad \text{Equation 2}$$

This type of process structure and probability resonates with the entropy measures used by previous authors; using derivations of entropy from the physical sciences or from information theory. For the purposes of this research entropy and information theory will, initially, be considered separately. The reason for this being that ultimately, the intention is to use the most appropriate application of the two when dealing with the structural and operational components of uncertainty discussed above.

2.3 Entropy

The origins of entropy can be found in statistical physics. Clausius formulated the second law of thermodynamics in circa 1865 via a statement that heat always flows from hot bodies to cold bodies; never the reverse. For this to happen matter must have an a priori state which he labelled entropy. The usual assertion is that entropy, from an information theory perspective, is a derivative of the physical approach. While the correlation between the two concepts almost holds, the correlation was discovered after the development of the information theory approach through the work of Hartley (1927) and then Shannon (1948), (Gliek, 2011).

The principle of entropy is that systems spontaneously move towards a state of disorder and confusion. To avoid any confusion, disorder and confusion in this sense are intended to describe a systems migration from a state of asymmetry, to a state of equilibrium at the observed level. To be clear then, order and clarity refer to states of asymmetry; disorder to states of equilibrium. In Information theory, entropy is described as the uncertainty of information sources during a communication process. In fact, Shannon’s (1948) argument was that if there was no entropy in the information process, then no information passed between source and recipient; an argument that aligns well with the point made by Williamson (1975) that business value is reliant on exchange over asymmetric states. Entropy then is the degree of uncertainty associated with the transition of information through the communication process. A simpler definition is proposed by Gliek (2011): ‘Entropy is the number of questions needed to arrive at the information required’.

Descriptions of entropy change with respect to the application. In information theory entropy is described as a numerical measure of the uncertainty of an outcome. In physics… a thermodynamic quantity representing the amount of energy in a system that is no longer available for doing mechanical work, entropy increases as matter and energy in the universe transitions to a state uniformity. In thermodynamics, entropy is commonly associated with the amount of order, disorder, and/or chaos in a thermodynamic system. In computing, entropy is the randomness collected by an operating system or application for use in cryptography or other uses that require random data. In this research the definition of entropy will be aligned with the information theory understanding; the origins of which – according to Gliek (2011) - seem to emanate from Hartley (1927) and Shannon (1948) at Bell laboratories where Hartley began to build the mathematical formula; beginning with:

$$H = n \log s. \qquad \text{Equation 3}$$

Where H equals the amount of information. ‘N equals the number of symbols in the message and s is the number of symbols in the language. The basic premise of the equation is that information redundancy is a necessary pre-requisite: the fewer symbols available, the more must be transmitted to get across a given amount of information (Gleik, 2011). Examples of this type of redundancy are easy to find, the police phonetic language is a good example.

The history of entropy as a measure in information theory, from its origins in the work of Hartley (1927) and Shannon (1948), would seem to be straight forward; this is not the case. Hartley's 1927 paper does not refer to entropy, neither does the formula in Equation 3 align with the popular perception of entropy as

$$s = \sum_{1}^{n} -p\, logp \qquad \text{Equation 4}$$

Furthermore, entropy in information theory has been questioned from a metaphorical and construction sense. There is considerable research already in place that uses Shannon entropy as a basis; however, because of the metaphorical and form construction issues, which will be detailed later, one of the intention in this research is to justify the use of entropy as a measure; this will include a review and critique of the work of Hartley and Shannon later in this chapter.

2.3.1 Entropy, and the Metaphor. Professor Ariah Ben-Naim in the preface to his 2008 book 'A Farewell to Entropy' quotes from Cooper (1968) who cites Clausius: 'I prefer going to ancient languages for the names of important scientific quantities, so that they mean the same thing in all living tongues. I propose, accordingly, to call S the entropy of a body, after the Greek word transformation. I have designedly coined the word entropy to be similar to energy, for these two quantities are so analogous in their physical significance that an analogy of denominations seems to be helpful'. Ben Naim (2009) continues with reference to Tribus's (1971) story on Shannon's naming:

'What's in a name? In the case of Shannon's measure the naming was not accidental. In 1961 Tribus asked Shannon what he had thought about when he had finally confirmed his famous measure. Shannon replied, 'My greatest concern was what to call it. I thought of calling it information, but the word was overly used, so I decided to call it uncertainty. When I discussed this with John von Neumann, he had a better idea. Von Neumann told me: you should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name. In the second place, and more important, no one knows what entropy really is, so in the debate you will always have the advantage'. Tsallis, on the first page of the preface to his 2009 book 'Introduction to non-extensive statistical mechanics' quotes the majority of the same text, this time referencing back to Tribus and McIrvine (1971) (readers interested in the further debate on this anecdote should refer to http://www.eoht.info/page/Neumann-Shannon+anecdote). Ben-Naim (2009) goes on to quote Denbigh (1981): 'In my view von Neumann did science a disservice. There are, of course, good mathematical reasons why information theory and statistical mechanics both require functions having the same formal structure. They have a common origin in probability theory, and they also need to satisfy certain common requirements such as additivity. Yet, this formal similarity does not imply that the functions necessarily signify or represent the same concept. The term 'entropy' had already been given a well-established physical meaning in thermodynamics, and it remains to be seen under what conditions, if any, thermodynamic entropy and information are mutually inconvertible'.

It seems, from this anecdote, that the metaphor applied to information theory was somewhat arbitrarily allocated on the basis that the formula looked the same and that there was already some confusion as to what entropy actually is. In addition, previous research on the use of entropy in operations and supply chain management has largely chosen to assume Shannon to be correct in using the term, and form, entropy; thus, these arguments have been developed from the Shannon assumption forward. This research intends to go beneath this assumption, to understand more of the function of the form in order to understand it's applicability in the business context used herein.

A second point is this, as we will discover later, entropy has been variously linked with degrees of information, uncertainty, mixed up ness etc. However, from an information theoretic perspective a simple description developed by both Gliek (2011) and Ben-Naim (2008) is this: It is the number of additional binary question needed to understand the state of the system. This is easily understood by example. Take 32 identical boxes and into one place an arbitrary object, mix the boxes up. It will take five binary questions to identify the box in which the arbitrary object resides (keep dividing the boxes in half). Another way of describing this is that it will take $log_2 32 = 5$ binary question. In addition, entropy is a function of volume; as volume increases so too does entropy, the thermodynamic term 'extensive' will explain this phenomena later in this section. Later, the text will demonstrate how this assertion misses a point on information structure, which calls for an additional dimension to be added to the assertion for it to remain valid; however, for now, the combination of these two points is the starting point for this research.

A further characteristic of the metaphor pertains to the use of probability and information. Entropy may, for now, reflect – as described above – the number of binary question required. However, this assertion does assumes the role of the experimenter to be the only role in the experiment. A fair coin has entropy S of $p \log_2 p = 1$. Once the coin is tossed entropy equals zero in the eyes of the experimenter and observers; however, to those who are unaware of the outcome – the information available from the experiment – the entropy - remains at one. This is further demonstrated in the classic thought experiment by Erwin Schrödinger: Outside the box the observer perceives a maximum entropy of one until the box is opened and the state of the cat known. Inside the box, the cat is pretty sure of its state throughout the experiment. The point is that to the observer not directly connected with the experimental activity, there is some form of hidden information that increases entropy. The concept of hidden information will be developed later.

Supply chains are getting ever larger, interconnected, mixed up, etc., generally for all the right reasons; to satisfy an ever more demanding customer. Organisation and governance structures are becoming increasingly complex in order to control the ever more complicated operations. As this volume of process and complexity increases, so to must entropy, being and extensive property (we are assuming the metaphor holds for the time being). Intuitively it seems that a measure of entropy would then provide an understanding of the state of the business that is not available using standard - macro level – measures or language; that is, entropy would provide a measure of the number of questions needed to understand the real state of the business. We will start by reviewing the origins of the term from a statistical mechanics and information perspective, I should point out that the intention is not to critique the formulation of the mathematical constructs except where the formulation is contributing to an understanding of the developing argument. The intention is to review the mathematical construct and the meaning of the term such that we have either a clearly defined understanding, or we have introduced a new term to fit the attributes and characteristics of the thing being measured.

It is important to clarify meanings applied through the rest of this thesis. Entropy will refer to the classic version of the form developed by Boltzmann (thermodynamics) and Shannon (Information theory). Other terms will be used throughout this thesis; for instance, ‘missing information’ or ‘hidden information’. Where these other terms are used, the relationship between the other term and entropy will be explained. Also, from a notation perspective, generally S is used to denote thermodynamic entropy and H for the information theory variant. Both of these notations will be used in this thesis to distinguish the type of entropy being discussed; however, practically, for the purposes of this research the two notations can be considered interchangeable. In some sections reference will be made to different forms of entropy, Renyi entropy, for instance; in these cases the form will be clearly stated.

The term entropy has been attached to the concept of diffusion or complexity as explained by Shannon. The metaphor has been attached because the formula for this diffusion or complexity closely follows the formula for entropy in the thermodynamics environment. This may be appropriate; that is, it is correct to attach the metaphor to the information theoretic formulation because the derivation of the formula in the information theory environment follows a similar logic to that in the thermodynamics environment, they are both based on probability theory. The purpose of the next section is to test this assumption and agree or disagree with the attachment of the metaphor to the common formula. The section is intended to explain the concept from an understanding perspective rather than a perspective of testing the proof for the formulaic development.

2.3.2 The Basics of a Classic Thermodynamic systems. The purpose of this section is to review the basics of the thermodynamic system; this will be done in two stages. Firstly, the text will review the basics of the problem of statistical mechanics; the understanding the state of the system, and the issue of the allocation of units of measure to the transition between the statistical version developed by Boltzmann, and the macro version developed by Clausius. This will be followed with a section to review the basic underpinning assumptions for the second law of thermodynamic that are required for entropy to be valid in this field.

2.3.2.1 Statistical Mechanics

Firstly, let us consider a thought experiment. Imaging two grains of sand, one blue and one yellow. Under a microscope, or with very good eyesight, they will be distinguishable and separable; that is, the state of the two grains of sand can be confirmed with some certainty. Now imagine 50 Kilos of blue sand and 50 Kilos of yellow sand both in separate bags. The state of each grain is difficult to analyse, but at least all the yellow and all blue sand is identifiable at the macro level. Now, imaging mixing the two bags together in a suitable container, 100kilos of green sand. At the macro level we now have this mixture that, at this level, appears green. The individual grains of sand have not changed colour, there are still approximately equal amounts of blue and yellow sand distinguishable at the micro level but indistinguishable at the macro level. So we now have 100 kilos of green sand, the construction or mixing of which was simple: empty into suitable container and stir. The deterministic laws of physics – Newton et.al – say that, in theory, processes are reversible. In theory we should be able to apply all the reverse forces we applied in mixing process and the grains of sand will proceed to un-mix themselves back into separate states. This is clearly not the case, there is something additional operating on the system that prevents or frustrates the un-mixing process such that it cannot occur. It was this type of problem that caused the great thinkers in thermodynamics (Clausius, Maxwell, Gibbs, Poincare etc.) to develop theories and rationales for the mixed up state of systems.

Let us now consider a different - but classic in the field of entropy - thought experiment that will enhance the journey towards our understanding of entropy. Consider a box, sealed off from its surroundings such that is nothing can permeate the sides of the box. The box is divided into two sections by a partition and in the partition is a very small hole. In one side of the box there are 1000 molecules of gas moving around freely and bouncing off each other and the sides of the box. Let's say that 99.9% of the time this bouncing around continues; but, 0.1% of the time one of the molecules passes through the hole into the opposite side of the box. Let's say that this happens every second so after one second there is one molecule, after two seconds two, three seconds three and so forth. If we take the state of the box after two seconds, with two molecules on one side of the box and 998 on the other. How many different ways are there for this to occur; that is, we cannot be sure which of the molecule have transitioned to the opposite side of the box, it can be any of the 1000. All we know is that 2 of the molecules have transitioned. Table 1shows how the number ways the probability of 1, 2, 3, etc. molecules increases for the first 10 occurrences.

Table 1: Table of molecule distributions.

Molecules on One Side	Molecules on the Other Site	Number of way of Occurring
1000	0	1
999	1	1000
998	2	499500
997	3	$1.6\ x\ 10^{8}$
996	4	$4.14\ x\ 10^{10}$
995	5	$8.25\ x\ 10^{12}$
994	6	$1.37\ x\ 10^{15}$
993	7	$1.94\ x\ 10^{17}$
992	8	$2.41\ x\ 10^{19}$
991	9	$2.66\ x\ 10^{21}$
990	10	$2.63\ x\ 10^{23}$

To put these numbers in perspective. At the point where there are 974 molecules one side of the box and 26 on the other, the number of ways in which this scenario can occur is

$$1.79\ x\ 10^{51} \qquad \text{Equation 5}$$

By comparison the estimated number of atoms in the whole of the world is suggested to be:

$$1.33\ x\ 10^{50} \qquad \text{Equation 6}$$

Considering only the first few combinations the numbers become vast; so let’s now repeat the thought experiment, but this time with a much smaller number of molecules; this time we will only consider 10 molecules. The same table, for only 10 molecules, is given in Table 2, and this time entropy has been included.

Table 2: Additional molecule distributions.

Molecules on One Side	Molecules on the Other Site	Number of way of Occurring	Entropy
10	0	1	0.003
9	1	10	0.020
8	2	45	0.060
7	3	120	0.110
6	4	210	0.141
5	5	252	0.150
4	6	210	0.141
3	7	120	0.110
2	8	45	0.060
1	9	10	0.020
0	10	1	0.003

We can see here that there is a distribution of probabilities. As the number of molecules on either side of the box moves towards equilibrium (5 on each side), the number of ways the scenario can occur increases to 252; that is, there are an increasing number of ways a particular scenario can occur as the probability of there being an equal number of molecules on each side of the box reaches equilibrium. So in this example there are $252/1024$ ways of achieving equilibrium and only 1 in 10 ways of having 9 molecules on one side, and one on the other. Equilibrium, or near to states of equilibrium are therefore more probable. Of course, both Table 1 and Table 2 are representation of $\binom{n}{k}$ binomial distributions where there are k successes in n trials and

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} \qquad \text{Equation 7}$$

Included in Table 2 is the Shannon/Boltzmann entropy; as we can see, as the molecules reach equilibrium, entropy maximizes; so in this experiment, entropy may be considered to be a measure of diffusion of the system. It is really important to note that this outcome does not occur as a consequence of the system being considered, it occurs as a feature of probabilities, in the same way that, for instance, sin θ occurs as a consequence of differing engineering scenarios (waves, pendulums and triangles serve as examples). This thought experiment serves to demonstrate the nebulous nature of the selection of the term ‘entropy’; which, basically and if we are to believe the earlier story, was selected because it represented the Greek word for transformation and because it was exactly this, nebulous according to von Neumann. We can argue that entropy is not a feature of statistical mechanics or information theory or, for that matter, any of the other applications referred to earlier in the text. Entropy is simply a feature of a distribution of probabilities. For example, the same result could be achieved by using marbles in a similarly constructed box, or elephants wandering around in two holding pens. In 1983 Jaynes (from Ben-Naim, 2008, Chapter 1, loc. 685.) stated:

‘The function H is called entropy, or better, information entropy of the distribution p_i. This is an unfortunate terminology which now seems impossible to correct. We must warn at the outset that the major occupational disease of this field is a persistent failure to distinguish between information entropy, which is a property of any probability distribution, and experimental entropy of thermodynamics, which is instead a property of a thermodynamic state as defined, for example, by such observed quantities as pressure, volume, temperature or magnetism of some physical system. They should never have been called the same name: the experimental entropy makes no reference to any probability distribution, and the information entropy makes no reference to thermodynamics. Many textbooks are fatally flawed by the author’s failure to distinguish between these entirely different things, and in consequence proving nonsense theorems’.

It seems the logic used for the application of the term ‘entropy’ needs not to be re-defined, but to be recalibrated: If we identified a particular business scenario that could be explained by, say standard deviation; following which a second, completely different business scenario was identified where standard deviation could again be used; we would not argue over the term standard deviation as applied to the two scenarios, instead we would seek to define the scenarios in which standard deviation can be applied. This seems to be the case with entropy: the formula $-\sum plogp$ can be applied to any probability distribution and should therefore be referred to as a statistical concept rather than an applied concept.

An alternative way of looking at the issue of terminology is through the formulations of the two statistical mechanics perspectives on entropy. Clausius’s original definition of entropy had units of Joules/Temperature, heat energy over temperature. Boltzmann’s definition $-\sum plogp$ is unit agnostics until Boltzmann’s constant k is added, at which point the unit of measure becomes Joules/Temperature and the value of the constant aligns the values of this form with those of the Clausius form. Without Boltzmann’s constant the measure is free of units. The purpose of Boltzmann’s constant is to correlate the outcome of the Boltzmann calculation with the Clausius calculation for the material being considered. Logically then, if we label only Clausius’s version as entropy, $-\sum plogp$ only becomes entropy after the application of Boltzmann’s constant. Prior to this application $-\sum plogp$ is unit less, nameless and a function of probability analysis.

2.3.2.2 The Second Law

The second law of thermodynamic was first stated by Clausius (c1850):

‘It is impossible to construct a system operating in a cycle which transfers heat from a cooler body to a hotter body without work being done on the system by the surroundings.’

Further attempts were made by Caratheodory (1909) and Plank (1927) to better clarify the second law and Li and Vitanyi (2008) more recently stated the second law as ‘No process is possible that has as its only results the transformation of heat into work’. It is important to recognise the second law – like the first - cannot be proven. It is only because the prediction of the second law are followed and the consequences of the law make sense, that the law itself is generally accepted; thus a corollary of the second law is that there exists a property of state, entropy S, defined as:

$$dS \geq \frac{dQ_{rev}}{T} \qquad \text{Equation 8}$$

Where dQ_{rev} is the reversible heat energy transfer to the system (Joules); T is the temperature in degrees Kelvin, with the equals sign indicating a reversible process and the greater than sign indicating an irreversible process. It is worth recalling the definitions of a system: Universe is totality and consists of a system in surroundings. Equation 9 refers to a system that does not reside isolated from surroundings or universe; thus, for a system to be reversible the entropy change of the system and that of the universe are zero:

$$ds(system) + \ ds\ (surroundings)$$
$$= ds\ (universe) = 0$$

Equation 9

For an irreversible process the total entropy change of the universe is greater than zero:

$$ds(system) + \ ds\ (surroundings)$$
$$= ds\ (universe) \leq 0$$

Equation 10

The classic view of entropy is further developed by the statistical formulation developed by Boltzmann and Gibbs. The basis of the statistical argument is that, at the macro level defined by the classic view, the state of the systems averages out at a given state or temperature or pressure. However, at the micro level, there remain differences in pressure, temperature, phase state etc. but this energy un-useable (Remember the mixed sand where the average colour is green). The classic explanation of statistical entropy is to consider – as was explained above - a finite number of molecules in a sealed box. In the centre of the box is a divider with a hole in it to let molecules pass from one side of the box to the other. The chances of finding a molecule in one side of the box or the other is $\frac{1}{2}$; the probability of finding all the molecules in one side or the other is $(\frac{1}{2})^n$ where n is the number of molecules. Boltzmann's argument was that there are statistically more probabilities of disordered states in a systems than there are of ordered states; thus, the macro state with the greatest number of accessible microstates is more likely; conversely, the macro state where all microstates are dynamically aligned, is the 'most improbable case conceivable'.

In statistical thermodynamics entropy s is defined as the natural logarithm of the number of distinct microstates w available, given the constraints of the macro system. Alternatively, entropy is a measure of the probabilities of the system distributed over the possible microstates.

$$s = klog_e w \qquad \text{Equation 11}$$

Where k is a constant of proportionality making the statistical form equal to classic version.

$$klog_e w = \frac{dQ_{rev}}{T} \qquad \text{Equation 12}$$

The assumed universality of Boltzmann – Gibbs classic definition of entropy has occasionally been brought into question. Tsallis (2009) quoted from Gibb's original research and suggested two assumptions for the applicability of the classic Boltzmann –Gibbs form:

1. The partition function Z must exist.
2. The form must be excluded from cases where gravity exists as an effect.

Some further explanation is required of these two assumptions. Firstly, the form in Equation 11, which incidentally, is carved on Boltzmann's gravestone in Vienna, is assumed applicable in cases where all cases of w are equiprobable.

$$p_i = \frac{1}{N}$$ Equation 13

For all w in cases where p_i is not equiprobable, the equation takes the form

$$s = k \sum_{1}^{n} p_i \, log p_i$$ Equation 14

This being the weighted sum of the probabilities. In this form p_i is the Gibbs distribution necessary for the above equation to function where:

$$p_i = \frac{e^{-E/kT}}{Z\ (T)}$$ Equation 15

And the partition function $Z\ (T)$ equals:

$$Z\ (T) = \sum_{1}^{n} e^{-E/kT}$$ Equation 16

Where E_i is the energy of the i th state of the system and T is the temperature. Thus, according to Gibbs, one of the originators of the form, entropy may not be a universal function and may only truly function if p_i follows the distribution explained above, and if gravity can be excluded. Hence, any use of the entropy form should perhaps address the Gibbs distribution question before its validity for any given application can be accepted. The problem of universality has no better illustration than that proposed by Maddox (1993) who explained that, for a black hole, entropy is not an extensive property, being a function of surface area rather than volume. The purpose of this research is not to prove, or otherwise, the thermodynamic validity of entropy; except insofar as, for rigour, the aim is to highlight the concerns in thermodynamic research over universality. Given this question over universality it is reasonable to offer modifications to the form and approach for applications, as has been the case with, for instance, the application of other entropy methods: Renyi, Kolmogorov, Kullbeck – Liebler etc.

This section has provided an overview of entropy in the thermodynamic sense. It has shown the Clausius, macrostate version of entropy, the Boltzmann-Gibbs microstate version of entropy, and the relation between the two. It has also explained the issues with the use of the term and how this impacts on, for instance, assumptions about the units of measure, the universal applicability and the prerequisites to the use of the Boltzmann - Gibbs form. It has already been stated that entropy has forms in thermodynamics and information theory, among others. We will now look at entropy from an information theory perspective and develop further an argument for a different approach to the use of the form.

2.4 Information Theory

As with entropy, and to hold to the research paradigm and method, this section will begin with an overview of the basics of information theory, before reviewing in detail the seminal works on the subject by Shannon (1948) and Hartley (1927). The section will then go on to review the existing research on the application of information theory in supply chain management, which includes the application of information theoretic entropy. Finally the section will look at the implications of the use of information theory in this research and propose a set of terminologies to be used moving forward.

2.4.1 The Basics of Information Theory. Language definition has a role to play in the definition of complexity and this research. Most of the definitions of supply chain management make some reference to the management of data or information; it therefore seems plausible that information theory could say something to the management of information within the supply chain. Take two binary strings of 30 digits:

101010101010101010101010101010
011011001011100110010001110110

Shannon (1948) questions the amount of information contained in each string. The more a string is compressible, the less information; in which case, the first string: print 10, repeat 15 times, stop, is a shorter definition than B and therefore carries less information. Kolmogorov complexity takes an algorithmic view of the same notion, asserting the complexity of an object to be proportional to its shortest description. Random, in this sense, refers to the degree of absence of regularity. The Kolmorgorov complexity of a system is defined as the computational resources required to specify the object (Li & Vitanyi, 2008). For instance, take to following string: abcabcabcabcabcabcabcabcabcabcabcabcabcabcabc; which could be described, using English language, as print 'abc' 15 times, stop; or, 15'abc', stop.

Or, take the following string: 1101001100; which has the property that position *z* is a 1 if the binary expansion of position z contains an odd number of 1's; that is:

Table 3: Binary expansion of position z.

Position	1	2	3	4	5	6	7	8	9	10
Binary Expansion								1	1	1
				1	1	1	1	0	0	0
		1	1	0	0	1	1	0	0	1
	1	0	1	0	1	0	1	0	1	0
Total	1	1	2	1	2	2	3	1	2	2
Thus	1	1	0	1	0	0	1	1	0	0

But what of the string ‘qpalzm56tygv’ to which there is a pattern that could be described. It is, obviously, more complex to do so and importantly, the length of the description or code that describes the string would be longer that the string itself. If this is the case, then the shortest possible description length for the string ‘qpalzm56tygv’ would be the string itself (Appendix E contains the pattern for the description of the string and more explanation on its derivation). Thus, the basics of information systems complexity are… For a given base language L, D is a description of the program that outputs a string z for a given input, including the length of any necessary integers. S is the length of description D. The number of characters in the description is the Kolmogorov complexity of D, written K(d). From a supply chain perspective, assuming a given language, the Kolmogorov Complexity of any integrated solution would be equal to the number of characters in the governed process model that consistently produce outputs for a given input. Put another way, the size of the governance documentation necessary to consistently be able to describe the operational state of the system is the Kolmogorov complexity. From this perspective, for any state space to be described (D), a language L will be required. A simple example serves to demonstrate this type of complexity: Let L equal the letters of the English alphabet, from which words are constructed. Let D equal the letters used to describe the term ‘supply chain management’. Using the definitions from section 3.1, Table 4 demonstrates the changing complexity K(d) of the definitions.

Table 4: Data count from supply chain management definitions.

Author	Year	Word Count	3 Period M/A	Character Count (inc. Spaces)	3 Period M/A Character Count
Oliver and Webber	1982	18		120	
Jones and Riley	1987	18		126	
Ellram	1991	19	18.33	111	119.00
Harland	1994	23	20.00	136	124.33
Berry	1994	38	26.67	279	175.33
Cooper	1997	17	26.00	110	175.00
Lee and Ng	1997	28	27.67	170	186.33
Handfield and Nichols	1999	54	33.00	360	213.33
Simchi and Levi	2000	49	43.67	330	286.67
Sterman	2000	20	41.00	109	266.33
Ayers	2001	21	30.00	131	190.00
Emmet and Crocker	2006	27	22.67	183	141.00
Wang	2007	22	23.33	159	157.67
Bowersox	2007	28	25.67	208	183.33
Sun and Ye	2008	29	26.33	208	191.67
Harrison and van Hoek	2008	32	29.67	182	199.33
Sokolov and Ivanov	2010	18	26.33	126	172.00

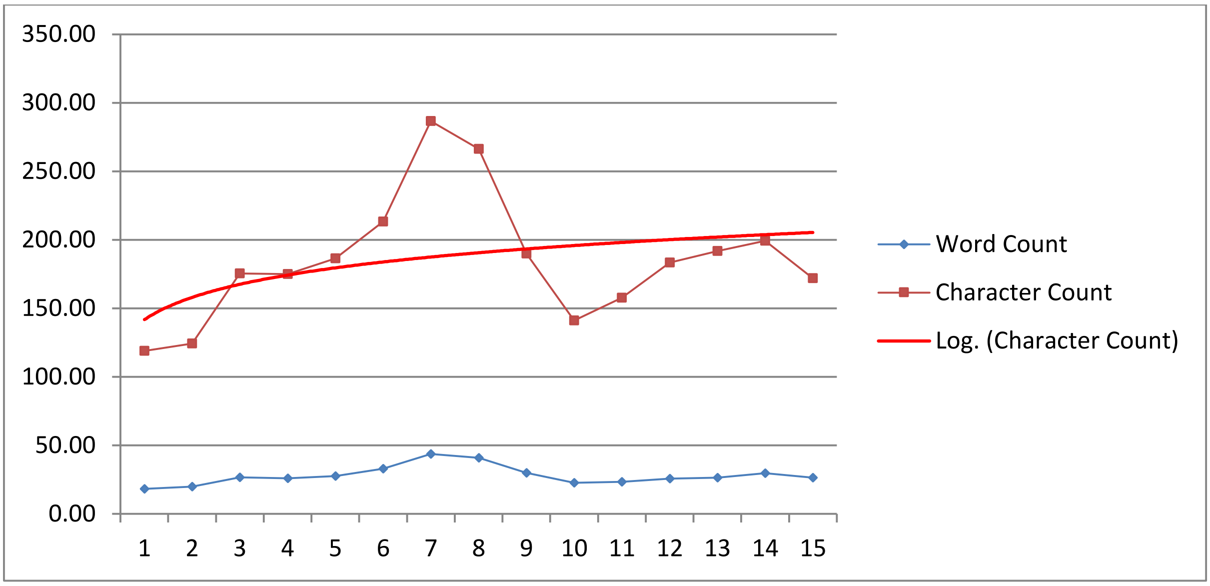


Figure 5: Graphical representation of supply chain definitions data count

In the supply chain management domain, the same approach can act to define the degree of complexity being managed by the organisation. Using an appropriate governance language, the number of processes necessary to define the scope and parameters of the organisations supply chain capability is equivalent to the Kolmogorov complexity (Kd) for the supply chain structure within that organisation. From an organisational perspective, a language (i.e. a governance structure) will be required to define the operating process scope and parameters (i.e. the state space of the business); from a supply chain management perspective, a common operating reference model would serve as this scoping and parameterising descriptive language. The supply chain operating reference model (SCOR), being one of a number an internationally recognised frameworks for the definition of supply chain processes could be used as such a language. Later in this research this framework will serve as a common language.

Supply chain management is obviously a multi organisational construct. Imagine two organisations A and B, each of which have a set of processes and parameters necessary to continuously describe the state of the system, to keep it in control, the governance system. The intersect (A;B) equates to the set of inter-related processes that enable one organisation to trade with the other (see Figure 6). For any organisation A,B,C…, ya, yb, yc, equals the set of governance descriptions necessary to describe the activities xa, xb, xc, necessary to complete the processes within the chosen scope of the business such that the business remains in control. Let D equal the description set such that D(y)= x. Or, y is the set of code words for the activity x with D as the decoding language – or simply, D(y) is the governance manual for process y that completes activity x. It follows then that if organisations are interacting in space A∪B, the same approach can be applied in the definition of the information theoretic complexity of the multi-organisation construction that defines a supply chain.

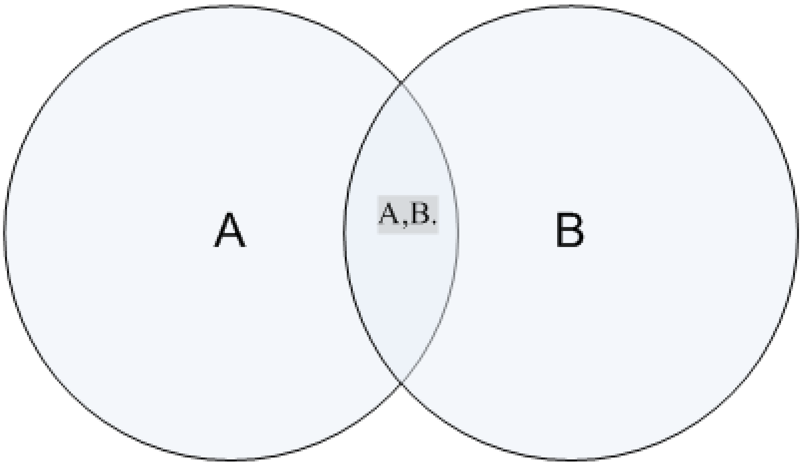


Figure 6: A,B and A∪B

Before going further, in order to develop an argument for the use of information theory as a contributor to understanding supply chain management complexity, a review and critique of the basis of information theory is required, built from the seminal works of Hartley (1927) and Shannon (1948).

2.4.2 A Critique and Review of the work of Hartley and Shannon. In his 1927 paper Hartley's aim was to establish a quantitative measure of the capacity of systems to transmit information. His meaning of the term information was specific and consisted of a group of symbols, for which there is a generally consistent meaning, that are, through some means, transferred between sender and recipient. Successive transmissions add to the meaning presented to the recipient.

Important, Hartley explains, is the understanding that an element of meaning is transferred through the removal of all other options available through the symbol structure. By this he means, for instance, the statement ‘the dog sat on the green mat’ can be interpreted as not any animal other than a dog; or, not a cat, mouse, hamster, elephant or some other such animal. Sat… did not lie or stand; on the not blue, yellow or red, not carpet or laminate. As a written explanation this is messy and complex, but the intent is to explain that the selection of a symbol serves to exclude all other symbols in the language. The selection of the letter A in the English alphabet acts to exclude letters B through to Z for the same location. For clarity we will refer to this principle as the Hartley exclusion principle.

Using morse code and the Baudot system as examples, Hartley then explains the need for, and differences between, primary and secondary symbols. If a primary symbol is limited to a number of states then a number of primary symbols can be combined to make a secondary symbol. The Baudot system uses clusters of five primary symbol binary states to allow for a thirty two character set of secondary symbols.

$$2^5 = 32$$ Equation 17

In more complex languages the numbers of options rises considerably. In the English alphabet, the potential number of three letter combinations is

$$26^3 = 17576 \qquad \text{Equation 18}$$

which is generalised as, the number of secondary symbols $s_2 = s^{n1}$ where n^1 equals the set size of primary symbols. Similarly, the number of secondary symbol sequences will be

$$s2^{n2} = s^{n1n2} \qquad \text{Equation 19}$$

Hartley suggests $n_{1,}n_2$ is equal to 'the number of primary symbols that would have been necessary to produce the same sequence had there been no mechanism for grouping the primary symbols into secondary symbols'. This can be used irrespective of groupings. Hartley then goes on to explain the measuring of the number of potential sequences. This is easily understood if we recognise that the purpose behind the selection of a symbol is to exclude all other meaningful symbols; thus he suggests an appropriate measure of information is the number of selections and not the number of potential symbols, a suggestion that will feature in the proposed measure herein. Furthermore, there is a need to balance the amount of information with the number of selections. To do so he introduces an arbitrary balancing value k such that $H = K_1 n_1 = K_2 n_2$; from which

$$H = K_1/\log s = K_2/\log s \qquad \text{Equation 20}$$

What Hartley has done here is to develop a measure of information within the parameters set out in his research, using the logarithm of the number of possible symbol sequences; for instance, in the Baudot case, the Hartley measure of information equates to:

$$H = nlogs = 5\ log2 = 5\ x\ 1 = 5$$ Equation 21

adding one to the number of possible sequence symbols gives…

$$H = nlogs = 6\ log2 = 6\ x\ 1 = 6$$ Equation 22

or 20% more information, which is demonstrative of the fact that the outcome from this form is volume dependent and therefore extensive -rather than intensive - in nature. With regard to form it is also worth noting that already there is a similarity between the Hartley form (Equation 8) and the Boltzmann form ($s = klog_e w$); this, even though Hartley does not use the term entropy in his paper.

In 1948 Shannon – from the same Bell laboratories as Hartley – published his work. The work had three main objectives: the effect of channel noise and, more importantly for this research, the savings available from the statistical structure and final destination of the information. Immediately, Shannon references the work of Hartley as providing the basis for his research.

Shannon breaks the concept of the communication, which we have to assume to be the transfer of information, into two aspects; the meaning of the message, and the systemic problem that a message is a message selected from a set of possible messages; the Hartley exclusion principle described above. Shannon also explains his reasoning for the use of a logarithmic content to his mathematics (Hartley does not justify the use of logarithms; the use emerges from the logic of the method) as:

- Practicality and alignment with other engineering principles,
- It is intuitively correct,
- It is a mathematically more suitable.

On page eight of his book on non-existential statistical mechanics Tsallis (2009) considers this justification to make available the idea of other entropies by quoting directly from Shannon: '*This theorem and its proof* ***are in no way necessary*** *for the present theory. It is given* ***chiefly to lend certain plausibility*** *to some of our later definitions.* ***The real justification of these definitions, however, will reside in their implications'*** (Tsallis added the bold typeface to highlight the point he was making).

Shannon also discusses the issue of logarithm base which affects the absolute values for the amount of information. He proposes the choice of the base for the logarithm should align with the number of stable positions for the devise; thus relays would be base two, having two stable positions. N such relays would therefore carry 2^N possible states.

In his discussion on discrete noiseless systems he represents the morse system graphically:

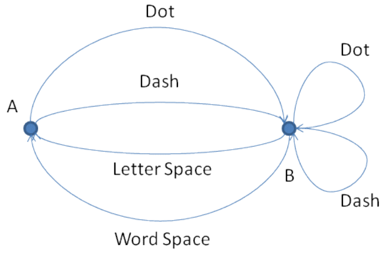


Figure 7: Shannon's network diagram for morse code.

For the purposes of this diagram the junction points represent the system state and the lines are the allowable symbols. While this represents the sentiment of the discussion, the grouping of system state changes to communicate symbols, it gives rise to two issues. Firstly, the states are unclear: A dot consists of 2t time units, one closed and one open (C,O), the starting state A must therefore be closed. Similarly as dash consisting of 4t units (C,C,C,O) starts from a closed state A and completes with an open state B. The letter space and word space consist of 3t and 6t units respectively all open (O,O,O) and (O,O,O,O,O,O) thus B, which is in an open state from either a dot or a dash, continues in the open state through the line or word space unit for 3t or 6t units respectively, at which point the state A will be open, contradicting the starting state for dot and dash communication and implying states A and B to be the same. In addition, given the state change during the process of communicating a dot or a dash, additional state spaces should be allowed. Secondly, the diagram does not describe the relationship of time to the state. The communication of the word space is three times longer than that of the dot. Finally, by comparison, Hartley is very clear on primary and secondary symbols and their relationship. Shannon is unclear on this point, referring to elementary symbols and denoting these with S irrespective of their primary or secondary status.

In his discussion on noiseless channels Shannon proposes a transmission rate measure given as N/second and explains the maximum possible transmission rate through one mode. This assumes that the state – 2 for relays, 26 for the English alphabet – is transmittable; that is, following Hartley, the transmission rate discussed by Shannon assumes the transmission state elements – the language – to be primary. Furthermore, there is a capacity measure issue that can be seen down the links. If a link has a capacity N(t) then the transmission capacity of the information necessary for the system to operate will be limited by this value. This phenomena can also can be seen in business models where link or node capacity is constrained by limiting factors. This can be demonstrated through for instance pert diagrams or adjacency or incidence matrices, hence the inclusion of network theory in this theoretical foundation. Network theory will be discussed shortly.

Shannon’s development of an argument of savings from statistical structure is based on the notion that the most popular state constructs should be communicated using the least number of symbols. According to Shannon, with reference to Chandrasekhar (1943), systems with the characteristics set out above, that generate symbols based on probability informed by preceding symbols, can be called stochastic. While the generation of symbols based on a priori symbols may be considered complex, the probability for a specific symbol, based on general probability and preceding symbols, can still be determined; the use of the term stochastic may not be wholly appropriate. Also, this does raise the question of the description to be applied to the scenario where there is equal probability and no dependency on preceding characters. In this latter case, the process of symbol selection would appear to be stochastic; thus, the term deterministic may be better used to describe the scenario referred to by Shannon. Whatever the label, the concept is important; it recognises that a probability can be applied to an outcome and an outcome that is informed by the preceding symbol, $p_i(j)$, where J represents the outcome given a probability p and prior state i; the structure discussed earlier. Shannon points out that the structure used in Figure 7 is a diagram; that is, the next state in the sequence is influenced by the preceding state for a given language. An extension of this approach is the scenario where the next letter is influenced by the previous two letters, the trigram, which is generalised as an n-gram structure. An example of the use of such structures can be seen in use in predictive text systems on mobile phones.

Shannon moves on to discuss *‘Choice, Uncertainty and Entropy’* where he begins by stating three axioms:

- H should be continuous in the p_i

- If all p_i are equal $p_i = \frac{1}{n}$, then H should be a monotonic increasing function of n.
- If a choice be broken down into two successive choices, the original H should be the weighted sum of the individual values of H. Figure 8 is taken from Shannon's 1948 paper:

Figure 8: Choice and successive choices (from Shannon, 1948)

$$H\left(\frac{1}{2}\cdot\frac{1}{3.}\cdot\frac{1}{6}\right) = H\left(\frac{1}{2}\cdot\frac{1}{2.}\right)+\frac{1}{2}H\left(\frac{1}{2}\cdot\frac{2}{3.}\right) \qquad \text{Equation 23}$$

Shannon's point is that H is common across both examples of 'choosing' and in line with point three above this is sometimes referred to as the 'independence on the grouping of events', which is a basis of another term 'missing information'. Missing information is defined as the amount of information one needs to acquire by asking questions (Ben - Naim, 2008) and it is assumed to be independent of the question structure. But in the above, Shannon fails to recognise the decision process assumption that supports the common position on H, and concurrently differentiates the position with regard to choices. In the left hand example above, a choice from three is made; that is, a question has been used that can distinguish between three outcomes. If this is the case, the number of questions needed will be

$$Q = log_3 3 = 1$$ Equation 24

By comparison, the right hand diagram, which does not change the value of H (the amount of entropy or unavailable information), does change the process of choosing, by changing the structure of the question asked from tertiary to binary, to understand the state of the system. Given binary questions, the average minimum number of questions required will be

$$Q = log_2 3 = 1.584$$ Equation 25

For the purposes of this research the omission of this structure of choosing or question setting is fundamental, and while uncertainty and entropy may be appropriate labels for discussion, it seems that the inclusion of choices confuses the issue and creates a requirement for much more clarity on choice, uncertainty and entropy. Choice, as increasing the understanding of the state of a system, is a function of the number of possible states and the question parameter (binary, tertiary etc.); clarifying or understanding this issue will form a significant part of this research.

Shannon asserts the form $H = -\sum p_i log p_i$ to be important to information theory as a measure of choice, uncertainty and information. We have already outlined above how choice, defined as understanding or selecting of a state, does not need the proposed form. As probabilities move towards equilibrium $p_i = \frac{1}{n}$ uncertainty also increases; i.e. maximum uncertainty and maximum entropy occur at the same point. We again need to be clear on terminology. In this case uncertainty is positioned as an adjunct to probability, the future is maximally uncertain because all outcomes are equally probable. However, to understand the current state of a system; that is, to reduce the observers uncertainty in the system, in scenarios where probability is in equilibrium or asymmetry, there is no reliance on probability of outcome, only a reliance on the number of question required to ascertain a state where each question has n possible outcomes.

Finally, from an information theory perspective, and re-iterating the basics of thermodynamics, Shannon's discounting of K (the thermodynamics equivalent of Boltzmann's constant) as '*the constant K merely amounts to the choice of units of measure'* creates a fundamental issue: What is the form H supposed to represent? Simply put, the form H without K is a numeric value derived as a function of probability, with no quantitative association to the topic under consideration. In its thermodynamic form K forms the relationship and conversion factor between the macro and micro variants of entropy, measured as Joules per degree Kelvin. To be meaningful, the form, from an information theoretic perspective, needs to have K to give the meaning to the value. Tsallis (2009) explains the same point slightly differently by distinguishing entropy and that which physically supports its notion; i.e. any probabilistic situation where $\sum p_i = 1$ can use the entropy form. Later on in his paper Shannon does state H to be measured in bits/symbol, this is simply too specific: Information, defined in bits, is only applicable for binary symbols or questions where $\log_2 n$ is applicable. The use of symbols or questions with greater than two outcomes would create a different unit of measure for H. For example, if there were a choice from ten symbols or choices then the outcome would be in dits ($\log_{10} n$). Furthermore, the right hand side of the H form does not include bits or symbols, it only includes probabilities and logarithms. The form H then needs an additional unit defining element to make the outcome, in Shannon's case bits per symbol, a valid measure. The consistent inclusion of K therefore seems mandatory for a unit of measure to be applicable to H. The form H does say something to information theory within the tight boundaries described by Shannon and Hartley; however, it seems that its use in the process of choice, and as an understanding of uncertainty may be too limited in meaning.

Hartley's research has largely been overlooked or superseded by Shannon's 1948 paper. This is, in the opinion of the author, unreasonable, the work of Hartley and Shannon are related but equally valuable with Hartley focussing on information as understanding and Shannon focussing on the diffusion of the information in the system. Given this research considers information and understanding more than it does diffusion of information it is likely that Hartley's research will be of greater significance as this thesis moves forward.

2.4.3 Entropy and Information Theory

2.4.3.1 Entropy at different levels. To understand the relationship between entropy and the approach proposed in this thesis there needs to be a clear picture developed on what entropy, in its Boltzmann - Gibbs form actually does. To do this a step by step approach to an example will be considered. In this example there are thirty two outcomes with nine possible states ranging from A to I.

Table 5: Thirty two events grouped into nine categorisations.

Outcome	State	State Count	P	Logp	p. logp
1	A	1	0.03125	-5	-0.15625
2	B	3	0.09375	-3.41504	-0.32016
3	B				
4	B				
5	C	4	0.125	-3	-0.375
6	C				
7	C				
8	C				
9	D	5	0.15625	-2.67807	-0.41845
10	D				
11	D				
12	D				
13	D				
14	E	6	0.1875	-2.41504	-0.45282
15	E				
16	E				
17	E				
18	E				
19	E				
20	F	5	0.15625	-2.67807	-0.41845
21	F				
22	F				
23	F				
24	F				
25	G	4	0.125	-3	-0.375
26	G				
27	G				
28	G				
29	H	3	0.09375	-3.41504	-0.32016
30	H				
31	H				
32	I	1	0.03125	-5	-0.15625

Stepping through this, P calculates the probability as the state count over the number of possible outcomes; for example, state E = 6/32 = 0.1875. $\log_2 p$ calculates the power by which 2 must be raised to equal p. So, for instance, for state A = 1, p=1/32 and 2 must be raised to the power -5 to equal 1/32. Similarly for state E=6, p = 6/32, 2 must be raised to the power -2.415 to equal p. To generalise:

$$p = \frac{\sum_{1}^{n_m} n}{N}$$ Equation 26

Note, as the probability increases, the log decreases. Another way to consider $logp$ is the number of binary questions required to find the state with probability p. State A has $p = 1/32$, thus it would take $\log_2 32 = 5$ binary questions to find the state A. Equally, it would take 2.415 binary questions to find any one of the six outcomes in state E. To be clear, 2.415 binary questions would find one only of the six E states, it would not find a specific state within the set of states E. To find a specific state would firstly require the identification of the remaining five events and then the identification of the specific state event. To identify each of the five remaining state events E would require an increasing number of binary questions ($\log_2 p$) as set out in Table 6:

Table 6: Binary questions for remaining state 'E'.

Remaining ‘E’ states	Remaining total	P	logp
5	31	0.16129	2.63227
4	30	0.13333	2.90689
3	29	0.10345	3.27302
2	28	0.07143	3.80735
1	27	0.03704	4.75489

The outcome of this process would be the identification of all six states E from the original group of thirty-two variables. To identify a specific state E from the group of six E states would then require a further group of questions.

$Plogp$ equals the number of binary questions for each given state and multiplies this by the probability that the given number questions will be required. For example, for state A, log 32 = 5 binary questions; there is p =1/32 chance that this number of questions will be required and consequently s = 1/32 x -5 = -0.15625. Similarly state E, log 6/32 = 2.415 binary questions for the state group multiplied by the group probability of 6/32 = -0.45282. To generalise, where there are N possible outcomes, $plogp$ takes the $x/N\,th$ fraction of the number of binary question required for any given x. S summarises $plogp$ for all N outcomes thus giving a weighted sum.

S, entropy in its Boltzmann - Gibbs form, then is the weighted sum of the number of binary questions required to identify a state group from the number of trials. An issue with this approach is created by the way the calculation evolves for given probabilities. To understand this in more detail a simple scenario will be used consisting of thirty two outcomes and only two states, A and B. In the first scenario thirty one outcomes are in state A and one in state B.

Table 7: 'A' primacy state count.

State	State Count	P	Logp	p. logp
A	31	0.96875	-0.0458	-0.04437
B	1	0.03125	-5	-0.15625

Thus S=0.2006

In the second scenario states A and B are equiprobable

Table 8: 'B' primacy state count.

State	**State**	**P**	**Logp**	**p.**
A	16	0.5	-1	-0.5
B	16	0.5	-1	-0.5

Thus S is maximally =1.

For completeness a third scenario where B =31 and A = 1 should be included, however, given the outcome would be the same as for scenario one, this scenario will not be included. In the first scenario, S being the weighted sum of the number of binary questions required to identify a state, multiplied by its probability equals 0.2006. Explaining this value using the approach described above seems challenging. S can be interpreted as the probable number of binary questions required to describe the state of the system, but in this scenario with a value of 0.2006, this is clearly not the case. Using binary questions would require $\log_2 32 = 5$ questions to identify the two groups; the value 0.2006 does not inform the question. The same assertion applies to the second scenario; dividing the thirty two outcomes in to two groups of sixteen is likely to provide two groups of sixteen outcomes each containing an equal distribution of states A and B. Generally then, entropy does not provide an indication of the possible number of binary questions required to identify given states. A statistical perspective does inform the degree of state diffusion; that is, the greater is the value of S, the greater the state diffusion. However, the measure is un-quantified being simply a product of probability.

Shannon, and other researchers applying this classic interpretation under the label of Shannon entropy, have utilised entropy as the amount of missing information. The implication in this usage is, continuing with the above scenarios, when all the thirty two outcomes are in either state A or state B, there is no missing information. When the outcomes are equally divided – as in scenario two – there is maximal missing information; i.e. generally, there is maximal missing information when state distribution across outcomes is equiprobable. By implication, information must be added to the system by transitioning the state distribution towards one outcome. Missing information then is the information that transitions the outcomes towards a common state. Entropy in this sense, does not inform the amount of missing information necessary to transition the business towards a specific state. It simply provides a measure of the 'mixed up ness' of the business states.

The above can be further understood by considering the 'state count' in the earlier scenarios. The state count equates to the number of occurrences of each particular state – in the same way that a total seven can be arrived at by throwing any of the six different combinations of two dice. Consequently, the state of the system is the sum of the states categorised as A or B; just like six is the sum of the states where two dice can sum to seven. Thus, as information is summarised into states, entropy reduces at the level of the summary; this is shown mathematically in Equation 29.

In business processes, while moving towards an outcome, the process is not missing information during that transition; it is a repository of information that goes through a set of process steps to a point where the repository is sufficient to trigger the next activity. Therefore, in business process terms, entropy provides an indication of the degree of mixed up ness of a business process at a particular point in the process and only at the level at which the process is being measured. This says little to the understanding the state of a system of industrial processes and an alternative should therefore be proposed that recognises the difference between understanding and ‘mixed up’, and recognises what seem to be horizontal and vertical dimensions – where entropy in the analysis of industry processes and structure have been considered; and the vertical dimension that creates some of the above issues.

An analysis of the proof for the form H in appendix A demonstrates consistency with the three assumptions proclaimed by Shannon (1948) and that n outcomes must be equal to H from their state groupings plus an average H within the state group events. Thus where state grouping exists - and to be consistent - H must consist of the H from the groupings plus and average H within the group events (H, information entropy and S, thermodynamic entropy, are interchangeable in this context). This leads to the fundamental issue with any assertion that entropy equates to a missing information. As will now be described, entropy is equal to missing information only at the level at which the analysis takes place hence; assuming a measure of entropy to equate to missing information without defining the level at which the information was derived is not a valid construct for the use of entropy as a measure of missing information. Re-visiting the data in Table 5, there are thirty two equiprobable outcomes, the outcomes have been categorised into 9 different state groups m with k elements in each. Entropy for the thirty two equiprobable outcomes equal:

$$S = \log_2 32 = 5 \qquad \text{Equation 27}$$

Whereas entropy for the grouped outcomes m equals:

$$S = \sum \mathrm{p_m} \log_2 p_m = 2.9925 \qquad \text{Equation 28}$$

The difference between the two values equates to:

$$S = {}^{\sum \frac{\mathrm{p_n}}{\mathrm{p_{m_k}}} \log_2 \frac{\mathrm{p_n}}{\mathrm{p_{m_k}}}}\Big/_{p_{m_k}} = 2.0075 = 5 - 2.9925 \qquad \text{Equation 29}$$

A diagrammatic representation of the phenomena described above is set out in Figure 8 and Figure 9:

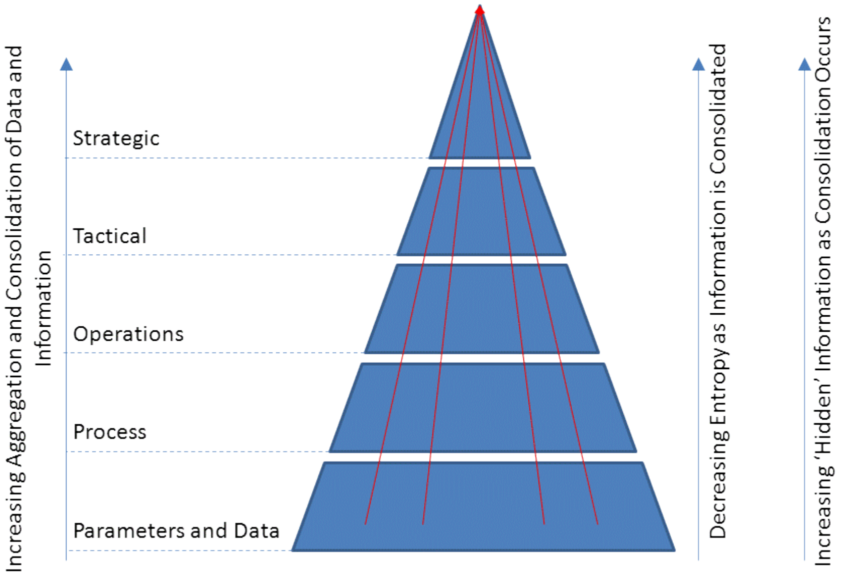


Figure 9: Diagram for the Effects on Entropy and 'Hidden Information' as consolidation occurs

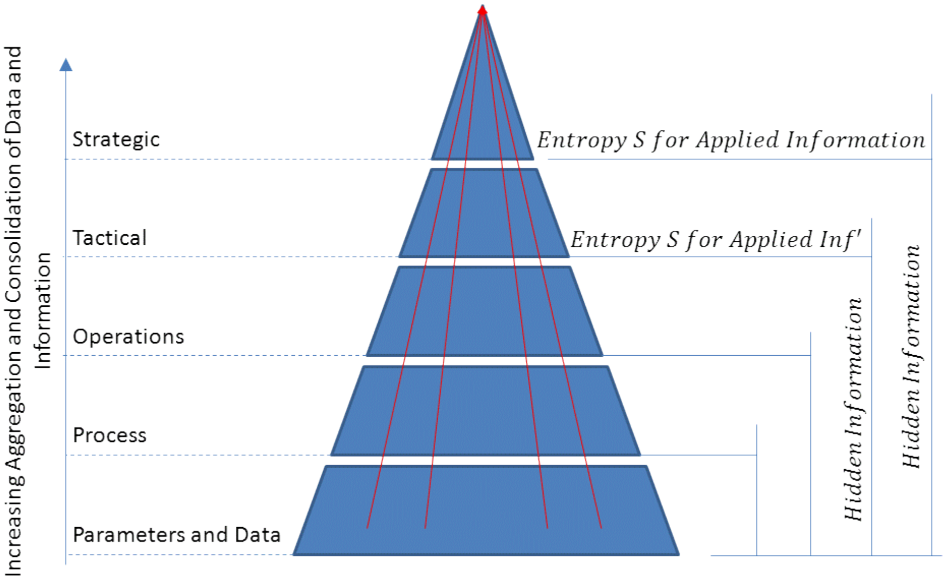


Figure 10: Diagram showing the dimension of Entropy and Hidden Information during the consolidation process.

We can draw an important point from these diagrams; that is, that using entropy as a measure is valid only at the level at which the data has been taken; therefore, when using a measure, it is important to select data at the level at which the measure is to be taken.

Additionally, it is important to understand the data level relative to the hierarchical structure of the data available.

2.4.3.2 Entropy as a predictive measure. Information theory, developed by Shannon (1948), as a process for applying probability to the selection of the symbols in a sequence, is aimed at ‘predicting’ the next symbol, his work had proved to be very valuable and can be seen in every day examples; no better an example of this is predictive text on a mobile phone. The approach assumes the purpose of the task to be to enable prediction, and in business processes the links between successive activities can have a predictable properties and measures. In a network structure, given any two outbound links with, for instance, a 60/40 outcome probability, entropy could be used as a measure of ‘mixed up ness’ for a given level of data. There are three issues with this approach: Firstly, as discussed earlier, entropy does not have a clearly defined unit of measure. Shannon, in his, 1948 paper suggested the base of the logarithm used, but the base of the logarithm is dictated by the number of possible symbols in the language, thus the language needs to be defined before a unit of measure can be determined. Secondly, from a management perspective, understanding the operational state of the business would not normally be based on prediction, it would be based on actual states. Thirdly, as we have just seen, entropy, as a measure and using the statistical form is only valid at the level in the business at which the data is captured. Measuring entropy at varying levels requires additional data. For these reasons, entropy, in its standard form is limited in its use. Never the less, entropy has variously been applied to the supply chain management domain and without addressing the fundamental issues highlighted above.

In summary, there are then three major issues with the use of entropy as a means of understanding the state of a system. It varies dependent on the level at which the measurement data is taken, the basis of entropy is probability, and the unit of measure is not definitive.

At this point clarity on metaphors is called for. Throughout the remainder of this thesis, Entropy will refer to the information theory variant of the form developed by Shannon (1948). The term ‘Missing Information’, sometimes used by authors as a description of the outcome of the use of the Shannon form, if used, will be specified at the time. The term used for what will become a proposed measure will be ‘Hidden Information’. The reasoning for this is that entropy, as a metaphor is tightly connected to the thermodynamic and information theoretic forms and the purpose of this thesis is not to argue over the correctness, or otherwise, of the connection – this has been done elsewhere; see, for instance Ben – Naim (2012) and Tsallis (2009). Secondly, the view proposed herein, is that to classify information as missing implies some sort of accidental omission: ‘you have missed this…’. While this may have been correct, the assumption made herein, and the approach that will be developed, is that the role of business management understands the structure of the business; as such, the aggregation or summarisation of data into information accepts that this process hides information. All that said, it is important to recognise how significantly and variously entropy had been used in business research and supply chain management.

2.4.4 The Application of Entropy to Supply Chain Management. The application of entropy to the business and supply chain domain is not new. Previous researchers have proposed the use of entropy as an analogous measure of the degree of flexibility in production routings (Yao, 1985). Extending its use, Kumar (1987) used entropy to measure loading and operational flexibility of a queuing network model. Rao and Gu (1994) continued to extend the use of entropy in this area by providing a quantitative measure of production volume and production flexibility. In communication theory entropy has also been used from an information theoretic perspective; most notably, Shannon (1948) who considered entropy to be the equivalent of uncertainty and, as supported by Shuiabi et al. (2004), as our perception of the world becomes increasingly complex, the number of phenomena about which we are uncertain increases and the uncertainty about each phenomena also increases. To decrease uncertainty one collects an ever increasing amount of information (Kapur & Kesavan, 1992. Sivadasan, 2002). Karberger and Mansoon (2000) stated two principles of entropy with regard to the economic system: Firstly, energy is neither consumed nor created in the economic process; secondly, every economic process results in an increase in total entropy. They go on to assert that technically, entropy is a measure of the distribution of energy among the degrees of freedom, with the degrees of freedom being a function of the composition of the system. This assertion correlates with the complexity view above, and with the discussion on Kolmogorov complexity and governance structures in that entropy is a measure of the distribution of effort among the structure of the system that comprise the degrees of freedom available in the governance structure. The mix up ness.

Of course, the question ‘why entropy?’ does arise; after all, what is wrong with the more traditional variance or standard deviation measure of dispersion. Dionisio et al., (2005), while investigating the use of entropy in financial markets, suggest that variance measures an average distance of outcomes of probability distribution from a given mean; whereas Ebrahimi et al., (1999) states that both measures reflect concentrations but their metrics are different; that is, variance measures concentration only around a given mean, whereas entropy measures diffuseness of density irrespective of the location of concentration, or, the ‘utility’ in the data (Dionisio et al., 2005). Ebrahimi et al., (1999) take the view that entropy, like variance, is an index; H is a measure of uncertainty of p. Entropy then, as a measure of dispersion or diffuseness, can be applied as a business measure of the process dispersion necessary to control the system, whereas mean and variance relate to the measure of volumetric throughput through the structure. For instance, for a given business, the mean number of customer orders received per day might be 100, plus or minus 20. The structural dispersion of the business function (the processes and parameters required to manage the customer orders) may have entropy of say 2.9. Increasing the mean increases resources with the same skills to process customer orders within the same parameters. An increase in entropy is driven by a greater dispersion of process and parameters, requiring different skills and processes. On this basis the two measures offer different measures of a business's operations.

Gonzales et al. (2010) recognised this difference but took a different approach by proposing three categories of business process measurement: understandability, complexity and reliability. They argue the case for effort to be applied to the understanding of these measures in two distinct business phases; the process design phase which is evidenced through the identification of the different measurable concepts for this phase, and the execution phase. In completing this work they identified entropy as one of the measures of business process understandability, complexity and reliability (See Table 9).

Table 9: From Gonzales et.al (2010). Understandability. Complexity and Reliability through Process Modelling and Execution phases.

Process Models	%	Process Execution	%
Complexity	44	Cycle Time	23
Understandability	21	Quality	11
Quality	7	Usability	11
Entropy	7	Functionality	11
Density	7	Cost	11
Cohesion	7	Effectiveness	11
Coupling	7		

Measures of complexity, in this sense, are aimed at understanding the effectivity of the structure and design of given business process model; hence the objectives of this thesis.

There is a further body of research, more specific to the application of entropy in the field of supply chain management, each of these will be reviewed in more detail. Blecker et al. (2005) developed an approach for analysing supply chain complexity. They argued the drivers of high complexity are customer tailoring, elaborate products, global procurement and distribution and that there is a direct link between efficiency and complexity; thus complexity management is a major issue for today's supply chains. They explain there are many varied approaches to, and definitions of, supply chains and they take the systems theoretic view of complexity. Referring to Luhmann (1980) they suggest complexity is determined by the amount and nature of the connections between the elements of the system, something that will be considered in more details as the argument in this thesis develops. They settle on a combination of Kersten's (2001) definition of complexity as variety in objects structures and processes; their own definition of internal and external components, and Frizelle and Woodcocks (1994) definition of structural and dynamic complexity components. Their approach breaks the system down into the standard systemic components: environment, system and subsystem, and they suggest a supply chain is a system comprising supplier, manufacturer and customer subsystems. Their approach develops a matrix where the attributes of their definition form the x axis and a derivation of Kersten's (2001) approach the y axis. At each intersect a list of characteristics is provided to represent the drivers of complexity of that type. The matrix is developed into a closed loop cause and effect model that encompasses three systemic levels. To capture information they present a detailed three tier framework; however, for the analysis of the data they suggest, rather than present, a variation on the method developed by Frizelle and Woodcock (1994). They conclude by suggesting their method can be used to comprehend complexity on three tiers. Their research is exploratory in that the framework presented is not tested out empirically.

Blecker et al. (2005) research looks at complexity as a structural component in the supply chain. Efstathiou, Calinescu and Blackburn (2002), on the other hand, using structural output in a manufacturing organisation, further develop the use of entropic measures to assess and monitor complexity. Specifically they use the output of the manufacturing schedule process, rather than the process of constructing the schedule, as input data. They follow Frizelle and Woodcock (1994) in defining structural complexity as the expected amount of information needed to describe the scheduled state of the system. Using the probability of resources being in allowable states, they apply Shannon's version of entropy to the probability for each resource and sum across all resources in the manufacturing facility. Thus, while the application of entropy is conceptually correct, the application will be limited by the issues identified earlier in this text; namely, units of measure, the fact that entropy is a feature of probability rather than a feature of a specific subject matter, and the lack of a value linking entropy to the subject matter and unit of measure (a K value). For dynamic complexity they use the same approach but this time using observed states rather than expected states; thus the same critique applies. They go on to develop a decision making complexity metric which they suggest satisfies six conditions:

- Decision-making complexity must increase with the number and types of parts and resources, and with the number of operations required for processing the part mix.
- Decision-making complexity must increase as the sequence flexibility for the parts in the production batch increases.
- Decision-making complexity must increase as the sharing of resources (either simultaneous or not) by parts increases.
- Decision-making complexity is dependent on the resource reliability and/or processing quality in a manner related to the way in which exceptional situations are dealt with, such as rework or processing a new part, when is the quality control made (after each operation or at the end of the processing), etc.
- Decision-making complexity must increase as the lot size decreases, as more decisions need to be taken.
- If the original part mix is split into two or more groups, then the overall decision-making complexity can be calculated as a function of the individual decision-making complexities.

Finally they conclude by describing how a web based expert system has been constructed; they summarise by confirming that the tool has been developed. There is no empirical data used.

Comparing Blecker et al. (2005) with Calinescu, et al. (2000) it seems that the term structure is interpreted in two ways. First, structure is considered to be the way the organisation is constructed, more specifically, the structure of the goverance process. Alternatively, structure is considered to be the output of the governance process, that which is planned to be completed.

Efstathiou et al. (2002) take the output from the governance process and categorise this as the structural component. Also, putting aside the criticisms regarding the use of entropy, the six satisfied conditions above readily connect with the construction of supply chains as a set of networked processes that become more complex with the increases in numbers and types of parts, production sequencing, decision making, and decreases in lot size. Viewing a supply chain as a network of inter-related processes will be developed from the next chapter onwards.

More recently, Isik (2010) takes globalisation, and particularly its effects on logistics as a key driver of supply chain complexity, as making understanding complexity in the supply chain a topic for research. Isik's approach is developed from Shannon (1948). The contribution of the work is to extend the formulas for structural and operational complexity and the aim of the research is to measure complexity associated with information and material flow, and demonstrate this with examples. Justifying the increase in complexity generally faced in supply chain management, Isik (2010) refers to increasing customer service for existing customers as a differentiator; increases in mergers and acquisitions, product launches and short product life cycles to protect organisations against the challenge of commoditisation and, for instance, increase the level of outsourcing to keep efficiency high and costs low. The inclusion of suppliers and customers in the issues faced by the supply chain suggests the challenges to be faced by the supply chain itself and not by the individual organisation therein. Isik also discounts any perception that complexity is simply a categorisation of either high or low by explaining that high complexity could be valuable in enabling a gain in market share; conversely low complexity might be expensive. Where complexity aligns with variety and where the customer is willing to pay for that variety, complexity is valuable. Whereas, for example, the effort required to manage variability of receipts and quantities on the supply side leads to non-value adding complexity and consequential costs. Understanding the value of complexity can therefore be beneficial to business differentiation.

Isik’s literature review covers all ‘usual’ papers (It includes Milgate (2001) which was criticised as being a plagiarism of Vachon, S. and Klassen, R. (2002)) and defines the elements of complexity that are of specific interest to his research as uncertainty and variability which he bounds - using the definitions proposed by Riess (1993) - as Size, Diversity, Variety and Uncertainty. The assumptions on structural complexity follow Efstathiou et al. (2002) and he develops the work of Frizelle and Woodcock (1994) by offering the view that complexity is not only a function of the probabilities of different states but that each state can have a level of complexity of its own. While not adding much detail to the point, this assertion aligns with the detail set out in 2.4.1; which was, where complexity is defined by binary question sets, there is, as is the case for thermodynamic entropy, a macro and micro component to complexity. Finally, Isik agrees with the selection of Shannon entropy but recognises that other entropies, namely Kolmogorov and Renyi, should be studied in other work. He concludes that the developed models work from examples but need testing and further research.

Sivadasan, et al. (2006) continue the themes set out in Blecker, Kersten, and Meyers (2005) and Efstathiou et al. (2002); this time analysing product quantity and time variation as variables. Complexity they define as the uncertainty associated with managing the variations in time or quantity across material or information flows at the supplier customer interface. They propose the use of an information theoretic mathematical model (entropy) and claim that the unique feature of this measure is that it measures, in relative terms, the amount of information required to describe the state of the system. Their intent is to demonstrate that the application of the measure will provide valuable insight in terms of degree of uncertainty, level of control, and detail of monitoring required at the interface. They characterise operational complexity at the customer supplier interface as visible via ad hoc orders, unreliable deliveries, changes to orders, alterations to specifications and other information changes.

In the same way that Efstathiou et al. (2002), build the idea that supply chains can be described as a network, Sivadasan et al. (2006) characterise the supply chain as interconnected, including these example aspects: number of elements or sub-systems, degree of order in the system, degree of interaction or connectivity, level of variety and degree of predictability. They follow Frizelle and Woodcock and differentiate between structural and dynamic complexity. Information theory is, they assert, a means of quantifying uncertainty, which they link to entropy by suggesting complexity needs increasing amounts of information to describe the state of the system which is what entropy is. Their review of information theory includes a comprehensive literature review of its use in manufacturing and non-manufacturing fields. They assert the information theoretic approach to be beneficial because it captures the ‘colloquial characteristics’ of complexity in one measure. Defining colloquialisms as every day rather than formal or literary may say something to this research in that herein one of the aims is to provide a framework that is colloquial and understandable in common language. They differentiate between fine and coarse grain measure, sticking with coarse grain because it allows for seven, plus or minus two, states they suggest to be the limits of human cognition; discounting fine graining as not of interest for day to day high level management. They conclude by suggesting structured analysis provides quantitative data for comparison across different areas of the business. The measure will change with the dynamics of the system. Not stated, but the implication is that line balancing by complexity becomes an option. They state the outcome of the research is valuable insight into degrees of uncertainty, levels of control and detailed monitoring requirements for the supplier customer interface. They also say complexity can be attributed to the number of elements and their interconnectedness.

The work of Isik (2010) and Sivadasan et al. (2006) thus further supports the view that complexity can be viewed as structural or dynamic complexity and, in these cases, they continue to relate structural complexity to the planned state of the business. The key addition from this research is the suggestion that there is an additional dimension to be considered: Isik describes this as each state in a complex situation having complexity of it's own, and Sivadasan et al. (2006) describe this as fine graining which they discount from their research. Both of these assertions align with the two dimensions of organisational structure and complexity set out earlier. These structures will be developed and analysed in more detail in chapter five.

Sivadasan, et al. (2002) contribute to the theoretical, conceptual and practical approaches to the use of information theoretic entropy based methodology for measuring and analysing operation complexity of supplier customer systems. They do this by providing an absorptive view; the extent to which organisations generate, absorb, export and import operational complexity. Their research provides a tool for identifying and classifying these four classes of operational complexity and they claim there was no measure of complexity in this context at the time of their research. They link complexity to firefighting internally, and with non-reliable relationships externally. Their definition of complexity is based on Frizelle's 1998 definition as variety and uncertainty associated with a system; from which structural/static complexity and operational/dynamic complexity are developed. Structural complexity is associated with variety embedded in the static system; dynamic complexity with the uncertainty in the external system. Their paper is limited to the development and presentation of a framework and method for measuring operational complexity where again, operational complexity is defined as the deviation from scheduled state. They create a mathematical description of their method and follow that with a conceptual framework and some simple examples; concluding the research to have provided valuable insight and enabled comparison across previously incomparable flows.

Gonzales, et al. (2010) provide a comprehensive systemic review of the literature on business process trends. They suggest there is little empirical testing in this field and suggest - in summary - the following key points have been observed:

- Complexity has been described as structural and dynamic (or operational). Structural complexity has been further differentiated as ‘structure as planned’ and ‘structure as governed’.
- The issue of fine graining, or micro analysis has been conceived but not analysed.
- Frameworks for this type of research have been attempted, but there is still a need for greater research in the general field.

More recently, Jung et al. (2011) take something of a lean approach to process uncertainty and variability, stating that the higher the complexity, the harder it is to understand or interpret the process model. Their research focuses on business process uncertainly and variability for which they propose the use of entropy based measures to quantify the uncertainty of business processes. They assert reducing uncertainty and variability in business processes leads to greater efficiency and predictability; conversely, they assert businesses with higher complexity have more difficulty with efficient planning and scheduling; however, they do not link this to customer service benefits as was the case in Isik (2010). They claim their measure enables a process which is easier to understand, less error prone, easier to manage and is more efficient;

the assumption being, as entropy reduces, the efficiency of the processes increases and the organisation becomes more lean. In applying entropy as a measure they recognise the issue of a constant K but chose to accept this as the value one. Their method assumes that business processes are linked sets of activities where the links are conditional on the decision made by the operator during an activity, giving rise to weighted links. Using historic data they determine probability; from this probability, using different examples, they calculate entropy asserting it to be a quantitative measure of uncertainty. They do not address the issue of entropy unit of measure; however, they go on to explain that a process of low entropy is, in the future, more easily predictable than one of high entropy. Thus, entropy in this context is an indicator of future predictability. They then go on to associate probabilities of a given state moving to connected states via links with probabilities. Similar to, but not described as network based, they go on to prove the mathematics for each type of split based on logical gates and use examples to justify their work.

There are a number of issues with this work: Firstly, the constant K is stated as ‘merely the choice of measurement unit’, however, they do not state what this unit might be. Without this qualification the values calculated cannot be said to be quantitative; they have no unit of measure. Secondly, they refer to the process models as petri-nets and describe them as bipartite graph structures. With the exception of one model, the processes can be classified as petri-nets but are unlikely to be classified as bipartite graph structure which required two interconnected sets. The third issue is with the process of business objects through an AND gate; from a governance perspective, dividing a business object to traverse two states concurrently is difficult to manage because the object must return to the origin to traverse the second ‘and’ activity. Usually, in such cases the business model is sequential, traversing one activity, then the second activity. Thus there is some question over the external validity of the AND gate scenario.

Rather than consider entropy to be useful as a measure related to uncertainty and variability Shuiabi et al. (2005) take the view that entropy may be a valuable measure of flexibility; specifically, the purpose of their paper is to ascertain entropy to be a measure of production volume and flexibility. Given the need for operational flexibility, there follows a need for a tool to measure such flexibility – they suggest. They provide a more thorough background on the use of entropy and its connection to the thermodynamic version. Specifically they propose entropy to be a measure of flexibility derived from the number of demands and the relative weight of each demand. They then constructed a simulation model through which a number of scenarios ran (increasing set up time, for instance). They concluded by confirming what was intuitively predicted and recognising the limitation of entropy as a relative and not absolute measure.

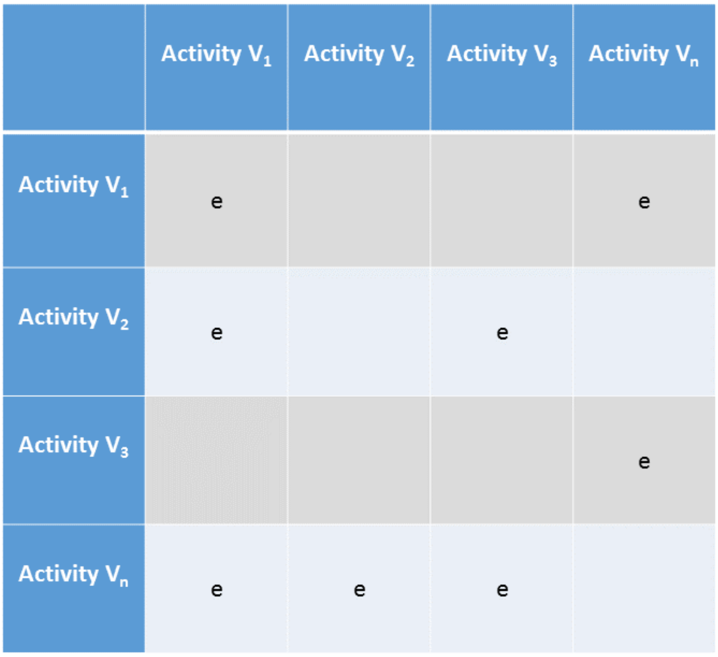
In summary, research in this domain has generally focussed on flexibility, diffusion of density, understand-ability, complexity, reliability, uncertainty and efficiency. Most of this research has assumed Shannon’s (1948) information theoretic approach to be a valid construct and has used the model as an input. As has been demonstrated, there are a number of issues with this approach, thus, this research aims to add to knowledge in this domain by adding an alternative construct that may be valid for use as a measure of complexity. Also, generally, research in this domain has followed some form of experimental approach, perhaps using field data but only as an input to the hypothesis. This could be seen to be reflective of the maturity of research in this domain and this research will continue in this direction in an attempt to further mature knowledge in the domain.

2.4.5 Implications for this Research. Entropy and information theory create a number of implications for this research. Firstly, entropy has both a horizontal and a vertical component. Sometimes these states are described as macro and micro states; sometimes as fine and course graining. The use of both components has previously been recognised but not studied; research is required to recognise and understand this two dimensional structure. Secondly, research on entropy in information theory has yet to provide a quantification for the values produced by the approach. Some form of applicable unit of measure is required. Thirdly, entropy as a metaphor is related to mixed up ness, missing information, diffuseness and other such descriptions. More clarity is required on what is being measured, through the application of a Boltzman constant equivalent, or some form of revised approach. Fourthly, Hartley's research is largely interpreted as a precursor to that conducted by Shannon who, with help – or possible interference - from von Neumann, connected the term entropy with information theory. The popularity and continued mysticism surrounding the term in such areas as cosmology has, to a degree, done Hartley's research a disservice by assuming his work to only serve the purpose of being a pre-cursor. There are elements of Hartley's research that continue to contribute or provide clarity in this research domain; for instance, through the application of his principle of exclusion or through a more appropriate use of metaphors for the addressed subjects (remember, Hartley did not use the term entropy). Consequently, this research will not lose sight of either authors' contribution to information theory research. Fifth, the use of entropy in research in supply chain management has largely ignored the difference between open and closed systems, and the application of entropy therein. This research will remain cognisant of this difference by applying entropy to only those elements of an organisation where a closed system is applicable; that is, the transactional activity of a business is an open system of interaction, whereas the governance activity of a business is a closed system of control. Sixth, research in this domain may offer new perspectives on the management of supply chain management; particularly as the pressures from globalisation, mass customisation and international cost arbitrage increase. This entropy/information theory perspective may offer a better colloquial understanding of complexity in the context of supply chain management; however, a framework and method is required to provide meaningful information, and a unit of measure for complexity is needed. An inter, and eventually intra, business area comparison of complexity units would be a unique and valuable insight into a business's supply chain operation; e.g. the equivalent of line balancing by complexity unit. Research in this area would also contribute to knowledge and provide valuable insight into the application of these theories in supply chain management. To do so a framework is required. The next section on Complex Networks begins to analyse the way in which such a network could be constructed.

2.5 Complex Networks

An argument for the analysis of supply chains using a graphical/network perspective is relatively easy to construct following the logic that a supply chain is a network, and the application of an appropriate network analysis method to the measurement of the supply chain as a graph provides valuable insight. This argument is valid but tautological: an analysis of a network is a good idea because the method of network analysis is a good idea. The purpose of this research results in an outcome that uses network or graph theory to construct a supply chain network suitable for the proposed analysis; however, the reasoning and logic developed for the use of network or graph theory is, as we have already seen, based on a logic that develops the need for the application of graph theory from the point of a business issue; in this case the complexity of supply chain management.

2.5.1 Complex Network Structures. We can conceptualization a supply chain structure as an n x n matrix where v, from the set V, is the set of activities necessary for the scope of the supply chain business and e, from the set E, are the links integrating the activities into a coherent structure. In network theory such matrices are labelled 'adjacency' matrices.



	Activity V_1	Activity V_2	Activity V_3	Activity V_n
Activity V_1	e			e
Activity V_2	e		e	
Activity V_3				e
Activity V_n	e	e	e	

Figure 11: The Adjacency matrix with connections e

If constructed in this way there are two distinct elements to the construction: First, there is the structure of the supply chain network as defined by the business governance model – the activities and connections are defined in the governance manual. Second, there is the structure of the supply chain as defined by the volume of objects (material, information, data, cash) passing through the governed structure. Earlier we saw how the structure of a supply chain represented as a matrix where the probabilities of each link being Pij (Xi, Xj) within the matrix and the purpose of this section is to further build on the build the network theory construct before integrating elements of the theory into a revised approach in the next section.

The representation of a supply chain as a network of interactions can be studied from the graph theory viewpoint; in fact, this adds another dimension to the current thinking on supply chain management as a set of interacting processes in that graph theory considers linkages as well as activities. Furthermore, the study of supply chain management as a network is essential: Dorogovtsev (2010) argues that specific network architectures, namely cage graph architectures, in synchronised systems - like a supply chain - offer an optimal architectural design. Before we consider supply chain management through this lens, an explanation of graph theory is required.

Graph theory studies the science of networks; considering networks as formed from sets of nodes or vertices and links or edges. In this section the terms nodes and links will be used. A graph G consists of a set of nodes n linked together by a set of links e; thus $G = (n, e)$. Links connecting nodes can be directed or undirected, and if directed they can be bi-directional or unidirectional. Bi-directional links are represented as arrows in a graph denoting two way linkage; unidirectional links being represented as arrows in a graph denoting a single direction. The total number of nodes n in a graph is v.

There are a number of different basic graph types, these should be explained as we move towards the representation of a supply chain network as a graph. A directed network is a network of nodes n from the set of all possible nodes N, connected by links e, from a set of all possible links E, where the links e are arrows signifying the unidirectional characteristic of the link. In a directed network the adjacency matrix - explained later - will be asymmetric about a diagonal; however, there is the case where a bi-directional link can be represented as two uni-directional links travelling in opposite directions. If this approach is selected for the representation of this network type, the corresponding adjacency matrix will be symmetrical.

The simplest form of graph is a tree graph. Tree graphs have specific characteristics: there is only ever one route between nodes on the tree, and the number of links can only be one less than the number of nodes; for a tree $e = n - 1$. For clarity, a number of basic tree graph structures are described:

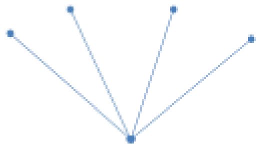


Figure 12: The star structure

The simplest graph of path form; maximum separation between nodes is two.

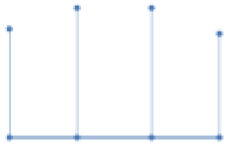


Figure 13: The comb structure

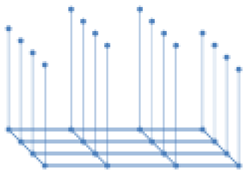


Figure 14: The brush structure

The lack of a loop in the set of tree graphs determines tree graphs also to be acyclic in nature.

Bipartite networks show links between separate sets of nodes *i* and *j* where links demonstrate connections between node *s* in one set, to nodes *d* in a separate set. A classic example of this type of network is the matrix of links between authors and papers referred to by Newman (2010), Dorogovsetz (2010) and Barret, Barthelemy and Vespignani (2011). (2011). In Figure 15 nodes *a* to *d* represent papers, nodes *i* to *k* authors. We can see how, in this scenario, authors are connected via shared papers; paper *a*, for instance, has had contributions by authors *i* and *j*.

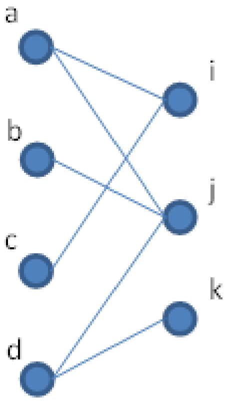
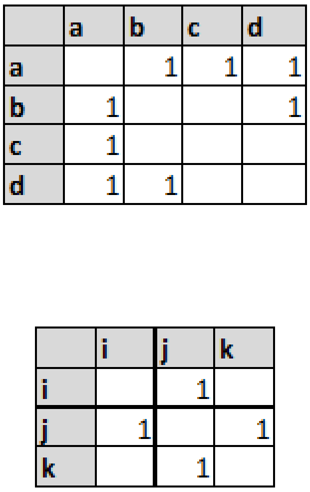


Figure 15: The Bipartite Graph

This type of form is referred to as a two mode projection (Newman, 2010). From these two mode projection, one mode projection can be derived. In the above example imaging an *i* to *k* matrix where the intersect is greater than zero if the authors contributed to a given paper, or if the matrix was defined *a* to *d* and the intersect represented a link of authorship. For Figure 15, the two matrices are shown in Figure 16.



	a	b	c	d
a		1	1	1
b	1			1
c	1			
d	1	1		

	i	j	k
i		1	
j	1		1
k		1	

Figure 16: Bipartite Adjacency Matrix (Incidence Matrix)

Newman (2010) differentiated further between two types of bipartite networks; the bibliographic and co-citation networks. Strictly speaking these two network constructs are different; essentially, the difference is that a bibliographic network recognizes the correlation between the bibliography of papers *v* in the set of papers *V*, whereas the co-citation network recognizes the correlation of authors.

Another method of describing the bipartite network is by using the hypergraph method; that is, when groups come together to form families with common characteristics, these can be represented as a hypergraph. A hypergraph model of Figure 16 is shown in Figure 17. The hypergraph model is more suited to social networks (Newman 2010); for this reason, if necessary, the bipartite representation will be adopted for this research.

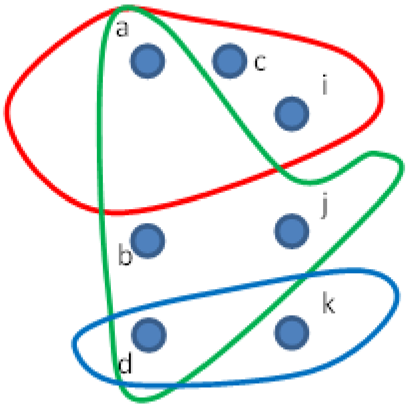


Figure 17: The Hypergraph Structure

Graphs can be considered to have two further categories; paths and loops. Path networks are those without loops as was the case with the tree networks. Path networks are acyclic whereas loop networks allow for cycles of connections in their structure. Below are three further examples of graph structures:

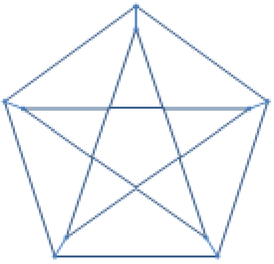


Figure 18: The cage graph

This, specifically, is a 3, 5 cage - or Petersen – graph; the main characteristics of which are: Each node has the same degree *k* throughout the graph (3 in this case), cage graphs have a minimum number of nodes for degree *k*. From a supply chain perspective graphs of this nature are important as in synchronised systems cage graphs enable optimal synchronisation (Dorogovsetz. 2010).

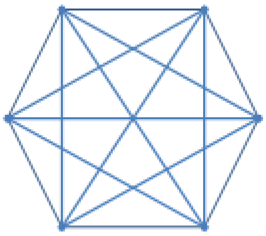


Figure 19: The complete graph

In the complete graph all nodes are connected to each other in a complete graph structure.

As explained earlier, links in graphs can be bi-directional or uni-directional. Another interpretation of this principle was offered by Newman (2010) who suggested pairs of nodes can be connected by more than one link. Graphs showing this characteristic are classified as non-simple multi-graphs. A further special case is where a single node links to itself; i.e. there is a feedback loop to the node. Nodes of this type are characterised as self-edging nodes. Graphs that do not have multi-edges or self-edging are classified as simple graphs. Figure 20 with Table 10 show a number of these different types of links.

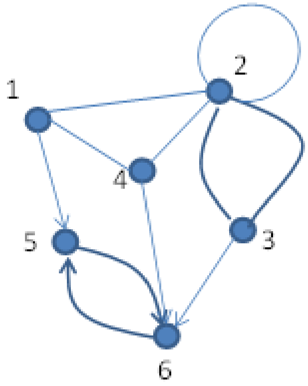


Figure 20: The Multigraph Structure

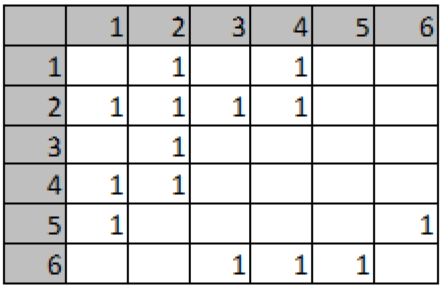
Consider the graph in Figure 20, consisting of six nodes and eleven links. The list of *e(ij)* links (1,2), (2,3),(2,3), (1,4),(1,5), (2,4), (5,6), (6,5), (4,6), (3,6), (2,2), represent a number of different link types, these have been described in Table 10:

Table 10: Table of Multigraph Link Types.

Link	Description.
1,2.	Undirected link.
2,3.	Undirected multi edge link.
1,4.	Undirected link.
1,5.	Directed link.
2,4.	Undirected link.
5,6.	Directed multi edge link
6,5.	Directed multi edge link.
4,6.	Directed link.
3,6.	Directed link.
2,2.	Undirected self edge link.

The concept of a graph as an adjacency matrix has already been proposed. The adjacency matrix for Figure 20 is shown in Table 11.

Table 11: The Multigraph Adjacency Matrix.

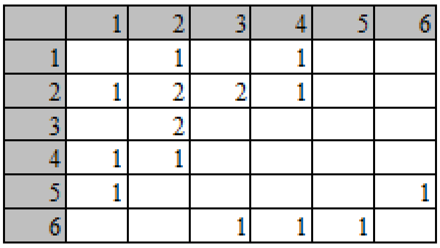


	1	2	3	4	5	6
1		1		1		
2	1	1	1	1		
3		1				
4	1	1				
5	1					1
6			1	1	1	

The adjacency matrix can be read as nodes *i* in column one outputting to nodes *j* in each row. So column four represents node four outputting to nodes one, two and six. Where a link is undirected, the adjacency matrix is symmetrical; for instance the link (1,4) is undirected and represented as a 1 at intersects (4,1) and (1,4). Directed links, (1,5) for instance, are represented asymmetrically with, in this case, the value 1 at intersect (1,5) and no value at intersect (5,1). The self-edge (2,2) can be seen at the intersect (2,2) in the graph, the only value on the diagonal. The diagonal of the adjacency matrix is represented as the intersects a_{ij} where ij = 1.1, 2.2, 3.3,..6.6. Values on the diagonal represent self-edge links.

To get a true representation of a graph a number of structural assumptions are required. In the case of undirected links both ends of the link will be represented by intersects a(ij) and a(ji); e.g. (1,2) and (2,1). Where a self-edge exists both ends of the link should recognized; in an undirected link this is done by adding the value 2 to the diagonal intersect, and 1 where the link is directed. Also, where undirected multi edge links exist, these too are shown with a number representing the actual number of links. Given this, a more correct version of the adjacency matrix is provided in Table 12:

Table 12: Corrected Multigraph Adjacency Matrix.



	1	2	3	4	5	6
1		1		1		
2	1	2	2	1		
3		2				
4	1	1				
5	1					1
6			1	1	1	

One final point on graph structure and the representation of such in adjacency matrices. So far the assumption has been a value in an adjacency matrix is used to represent a yes/no state: 1, the link exists; 0, no link exists. Imagine a scenario where the link e is given a value for, say, the volume of data travelling along the link; so, for each link and node a 'weight' is applied to represent the volume, or some other such characteristic, that traverses the link. This type of matrix is referred to as a weighted graph. Another way of considering this is that links in a given graph have a standard value; in which case the number of links will represent the correct number of links between given nodes to represent the weight of the connection. In these cases the weight of the link in the graph and associated adjacency matrix is the given number of links necessary for the defined graph structure. The use of these weighted graph structures is important when considering the dynamics of the material and information traversing the supply chain governance structure. For instance, if two nodes represent two logistics hubs and the link is representative of the vehicle used to link the two hubs. The weight of the link could be equal to the load capacity of the vehicle, and the number of links representative of the number of vehicles necessary to transport all the materials between the two logistics hubs. The ability to analyze a network or graph is important to the later elements of this research; however, for now, a basic explanation of the analytical concepts is provided:

Graph density is the degree to which the graph is connected and is defined as

$$D = E/(\frac{N(N-1)}{2}) \qquad \text{Equation 30}$$

where D = graph density, E is the total number of links in the graph and N is the total number of nodes in the graph. The lower the density, the greater the number of zeros in the corresponding adjacency matrix. Degree k is defined as the number of links incident on a specific node. Degree is therefore a local characteristic of the graph. For undirected graphs degree k will be symmetrical; that is, for each x_{ij} there will be a corresponding x_{ji} and the adjacency matrix will be symetrical. For directed graphs this assumption of symmetry cannot be made: k_{in} will be the inbound degree for a specific node; k_{out} will be the outbound degree; but, k_{in} should equal k_{out}.

$$kin_i = \sum_{j=1}^{n} (x_{ji}) \qquad kout_j = \sum_{i=1}^{n} (x_{ij}) \qquad \text{Equation 31}$$

Links in real networks can be strong or weaker; for instance, in a make to stock environment, incoming customer orders will largely follow a path of acceptance, pick from stock, pack and ship. Thus the governance route selected will be probabilistically determined based on a set of predetermined parameters and follow some form of probability distribution through the governance structure. Similarly nodes can be seen to be relatively more important or heavier in the graph scheme. In addition to a method of structuring a graph through links, nodes and a representative adjacency matrix, a method for representing the characteristics of relative strength is required to make a graph representative of the real work. To do this, two additional characteristics can be defined. The strength of a link is measured using a real world measure of the relative strength of a particular link; for instance, a specific link can be measured in terms of traffic volume that occurs across the link between two nodes. Similarly the weight of a node can be measured by the sum of the weights of the k degrees. This too can be represented in the adjacency matrix. To be clear, the example of a hub and vehicle is used to explain one possibility for defining the weight of a link and the number of links. The point here is different: using the in degree or out degree of a node, group of nodes, or the whole matrix, provides some insight to the connectedness or density of the structure.

2.5.2 Basic Random Network Models. Dorogovtsev (2010) defines a random graph as a statistical ensemble whose members are all possible labeled graphs of given number of nodes N and links L were all these members have equal statistical weight; that is, within the boundary of their construction, they are structurally homogeneous. The two principle network models are described as the Gilbert, and Erdos-Renyi, (denoted G_{np} and G_{ne} respectively) models after their respective originators. There is a difference in the definition for these two basic models. The Erdos Renyi construct begins with the set of N nodes which are linked at random by E links. In contrast, the Gilbert model begins with the N nodes and with all possible $N(N-1)/2$ links being assigned a probability. The relationship between the two can be defined as:

$$P(G_{ne}) = P^{E}(1-P)^{\frac{1}{2}N(N-1)-E} \qquad \text{Equation 32}$$

This last point is an issue for this research. In a real supply chain a graph representing the governed structure of the supply chain is unlikely to have all links such that the graph is complete; neither is the graph likely to be linked randomly; but in contrast the distribution of processes or the connectedness of each node does say something of the organisational structure necessary for the effective management of the business. To further explain: If the nodes of a graph represent business activities and links represent the connections between these activities. At points on the graph it is likely that decisions taken in an activity lead to a connection with more than one alternative activity as an outcome. If this is the case, the ratio of activities travelling down the link options would be represented by the relative strength of each link. Similarly, the average weight of a node, in business terms, represents the average unit volume transitioning a specific set of capabilities in the business. This volume measure of activity assumes a structure to be in place on which the volumetric activity takes place; it says nothing to the total range of activities required by a business, nor the diffused set of parameters required to be managed by the business to ensure the correct functioning of the volume activities. The relationship between Gilbert and Erdos-Renyi structures, and the range of activities necessary to operate a supply chain can be understood respectively as: The Gilbert random graph begins with nodes n linked to all other nodes n-1 by links with assigned probability. The supply chain network, represented as a similar structure, would see each activity in the supply chain connected to every other activity in the supply chain by a link with an assigned probability, with the probability being zero where the connection was illogical in the governance process. Similarly, with the Erdos - Renyi structure, the supply chain network will be represented as a set of links *e* from all possible sets of links *E* such that the logic of the supply chain governance structure holds true. Hence, from the random graph perspective, the governance process defined by the actual structure of the supply chain network is a specific case from all possible cases in either the Gilbert or Erdos-Renyi random graph approaches.

Earlier it was explained how the use of information theory and specifically Kolmogorov complexity (To recap. The basics of Kolmogorov complexity are: given a base language *L*, *D* is a description of the program that outputs a string z for a given input, including the length of any necessary integers. *S* is the length of description *D*. The number of characters in the description is the Kolmogorov complexity of *D*, written *K(d)),* from a supply chain perspective and for any integrated solution will be equivalent to the number of characters in the governed process that consistently outputs string *z* for a given input and base language. From this perspective, for a generalized approach, it is reasonable to suggest that for a given state space to be defined, a language will be required. A simple example serves to demonstrate this type of complexity: Let *L* equal the letters of the English alphabet, from which words are constructed. Let *D* equal the letters used in a definition of the term ‘supply chain management’. The length of the description *S* = 23 and the number of characters in the description *K(d*)= 15. The correlation between the information theory approach and the network analysis approach is immediately visible. Base language *L* aligns with either the Gibert or Erdos -Renyi models of statistical ensembles of networks through the representation of the set of all possible elements of the language in the same way that the Gilbert or Erdos_Renyi models represent the ensemble of all possible models. In information theory, *D* is the description of the programme, *D* being a sub set of *L*, the output of which is *z*, which aligns with the specific Gilbert or Erdos-Renyi model, selected from the statistical ensemble of models that satisfy the supply chain governance process of the business. Conceptually it follows that the output of the network, selected from an ensemble of networks is a governed process that outputs *z*, which is the Kolmogorov complexity *(Kd)* of the network; i.e. language *D* is the governance of the network that produces output I.

Before we go further with the use of graph theory, some history and a further basic definition needs to be set out. Leonard Euler worked, except for a considerable break in Berlin, at St Petersburg University from the time he was 20 until his death in 1783 at the age of 76. In 1735 he was invited to study the Konigsburg bridge problem, which is now seen as the origin of network theory. The problem, at face value, appeared simple; four land masses connected by 7 bridges and a question: Was there a path available where a person can cross all the bridges only once? Euler proved such a challenge to be impossible to complete. Ninety years after Euler's death, Carl Hierholzer proved that such a walk is only possible if each node in the graph has an even number of links.

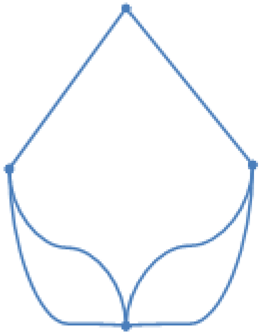


Figure 21: The four nodes and seven links of the Konigsburg bridge problem.

Given this explanation of network theory a number of basic assumptions can be made. Firstly, for a set of random graphs, the average number of links will be:

$$[E] = \frac{N(N-1)p.}{2} \qquad \text{Equation 33}$$

Thus the average degree can be deduced as:

$$[K] = \frac{2E.}{N} \qquad \text{Equation 34}$$

These two measures imply a degree of equilibrium across the network. There is empirical evidence (Barret et al, 2011) to suggest that, in some cases, this is not the case and that nodes tend to link to nodes with similar properties; this is typically visible in social sciences and epidemiology research. ‘Assortative mixing’ is the term used to refers to this tendency to link to nodes with similar characteristics. Conversely, the term ‘disassortative mixing’ refers to nodes with a tendency to link to nodes with dissimilar characteristics.

Clustering, referred to as transitivity in the realms of sociology (Wasserman & Faust, 1994), refers to the tendency in graphs for 'cliques' to develop and can be explained, in an undirected graph, as the tendency for nodes linked to node i to themselves be linked. If node i is of degree k, and these nodes have links e then

$$C_i = \frac{e}{k(k-1)/2} \qquad \text{Equation 35}$$

and

$$e_i = \frac{1}{2} \sum_{jl} x_{ij}, x_{jl}, x_{li} \qquad \text{Equation 36}$$

And the average clustering co-efficient is given by

$$\langle C \rangle = \frac{1}{N} \sum_{i} C(i) \qquad \text{Equation 37}$$

Clustering is only meaningful for degree $k_i \ > 1$.

The clustering effect gives rise to communities or graph components, the nodes of which may or may not, be reachable from other nodes in the graph. A network can include one or many components and a component is defined as a network subgroup where the nodes are connected to each other by at least one link and where this logic cannot be held with the addition of another node. The idea is simple for undirected graphs; but in the context directed graphs, the idea is more complex. To simplify the understanding of directed components, three descriptions are generally provided. Firstly strongly connected components are used to describe a situation where there is a path *AB* and a corresponding path *BA*; thus, acyclic directed components cannot be described as strongly connected. Secondly, an out-component is a directed network the nodes of which are reachable from the origin (i.e. you can get out from the origin to the component, but you cannot get from the component to the origin). Finally, an in component graph is a directed network the origin of which is reachable from the nodes (i.e. you can get from the graph component to the origin, but not from the origin to the component).

The major works on network theory consider it as a structure based on the above measures and constructs (Dorogovtsev, 2010; Newman, 2010; Barrat et al. 2011). But there is a need to consider a network as a layered construct and the work of Easely and Klienberg (2010) stands out in this respect. They discuss the use of breadth first analysis, a method that considers the network as a stratified construct. The stratified construct is how governance constructs in supply chain management are structured; hence the construct in this research will consider the supply chain to be a layered structure.

A supply chain consists of a set of activities – nodes - connected by a set of links. The supply chain activities can sit across multiple organizations or inter-organization business units with clear input/output methods between each. A purchase order output and customer order input transactions between organization A and organization B would be representative of the link between two giant components or communities of interlinked processes within a supply chain network; which could be represented in the format set out above. The use of graph theory in the analysis of supply chains has received some attention in the academic literature; the following section outlines the salient points in this research.

2.5.3 The Application of Graph Theory to Supply Chain Management. Smith (2012) highlights the point that there is a timescale issue when thinking about topology and dynamics; the point is that the timescale between the development of the topology and the effective dynamics can be long. The implication is that the topology may well be out of date or in need of modification, or at least may be incongruent with the requirement of the dynamic. Smith's application of graph theory is used to point out the potential of a time lag between the construction of the structure and the impact this structure has on the dynamic activity in the network. Through his research he is keen to stress the difference between a bullwhip effect and trophic cascade affect; his argument being that while the characteristics may look similar, the basic underlying problem is different, and thus any given solution may also be different. Anand

and Bianconi (2009) used Gibbs entropy as a measure of complexity in the Petersen graph structure; accepting an assumption of a fixed number of links per node; thus allowing Gibbs entropy to be in a simple form:

$$S = 1/N \, Log \, N. \qquad \text{Equation 38}$$

Where all $N's$ are equiprobable in the analysis. Anand and Bianconi also point out, from a complex network perspective, and referring to the work of Fortunato (2009) and Gfeller, Chapellier and de los Rios (2005), that there is a gap in the research on the state and content of information in a complex network; hence the intention of this research is to go some way to closing that gap with specific reference to supply chains as complex networks. Sole and Valverde (2004) take an alternative perspective by suggesting complex networks to be heterogeneous in nature, with the probability distribution of degree K being a useful measure of heterogeneity. More specifically, their interest is on using degree distribution in the assortative and dissortative clustering scenarios; hence they use remaining degree distribution as a measurement basis. They go on to use entropy as a potential measure of average heterogeneity for a given network where their approach to graph construction combines the basic principles of the G_{NE} and G_{NL} approach, in that the probability is assigned to a given set of links, specific to a given graph. Thus, this type of mapping would be specific to each individual supply chain: From all possible activities (the language *D*) select those necessary to govern the business process (language *d*); link through process mapping to determine a degree distribution for a given structure.

Lassen and van der Aalst (2009), from a basis of business process modelling languages initially state, following but without referring to, the principles of Kolmogorov complexity, the complexity of a process model to be a reflection of the problem to be solved. They support this assertion using empirical data from previous research which demonstrates that twenty percent of an event driven process chain model had errors. Also, it could be that a given model is unnecessarily too complex - evidence in their research concludes processes can be overly complex which will lead to additional costs, a point also made by Isik (2010) – and therefore it seems logical to be able to understand the complexity embedded in the process models; thus, they assert, the complexity of a model impacts on the quality of the solution. They present three complexity metrics with a focus on understanding 'structured- ness'. The first two metrics are extensions of existing metrics; the third metric '*is a new metric that better tries to capture the complexity of a model as it is perceived by humans'.* They differentiate between the 'static' and 'dynamic' models, using 'petri-net' and 'workflow-net' respectively; where a petri-net is representative of directed bipartite graphs. Petri-nets form the structure of the network and workflow nets form the object or case model though the Petri net. As is often the case in this research field, they conclude that their revised metric contributes to the theory and offers some insight, but further field testing is required to validate the results.

Battini, Persona and Allensini (2007) combined a network and ecology view in proposing a measure of supply chain complexity. By analogy they assert similarities of network systems, flow and nodes between eco-systems and supply chains and thus propose a measure as a development of the works of Frizelle and Woodcock (1994), Calinescu (1998), Sivadasan (2002), Meyer and Foley Curley (1995), Efstathiou (2002) and Arteta and Giachetti (2004) by linking this work on supply chain complexity with the work of Ulanowicz (1984, 2003 and 2004). Their method is to apply a graph theory approach to the construction of a weighted network for a given supply chain and develop two measures: Firstly a total system throughput measure representing the size of the system and equating to the sum of the weights of the links in the network; and the average mutual information which uses a variation of Shannon conditional entropy for weighted in degree and out degree measures in which average mutual information decreases and entropy increases. The application of the total system throughput and average mutual information formulae are then embedded in a seven step approach and applied, using largely estimated data, to a transaction cost model. They conclude by acknowledging this to be a first work on the application of ecology systems theory to supply chain networks but suggest the approach can be seen to be valuable and has further value in the field.

The above research develops the use of networks analysis to the field of supply chain management. Other research, not specific to the application of network theory to supply chain management provides some further insight into how complexity may be considered in the field of network theory.

2.5.4 Alternative approaches to studying Network Complexity. Sanchez - Gonzales, Garcia, Ruiz, and Mendling, (2012), concurring with the business process quality issue highlighted by Lassen and van der Alst (2009), took a quality perspective to measuring process models by measuring process gateway parameters. Their method was to use a controlled experiment to evaluate uncertainty and modifiability as process quality characteristics. Six gateway parameters are considered, defined as types XOR, OR and AND combined with the directional characteristic SPLITS and JOINS. Six complexity measures are considered: Control Flow Complexity (CFC), Gateway Mismatch (GM), Gateway Heterogeneity (GH), Average Gateway Degree (AGD), Maximum Gateway Degree (MGD) and Total Number of Gateways (TNG). In studying these characteristics their research question asked if it is possible to automatically distinguish between understandable/modifiable models and those not so, through measures of structural complexity using the above metrics. They concluded that threshold values for each of the above measures were valuable in the process of understanding and modifying business processes. These threshold values are set out in the following points:

- In a business process, include no more that 18 to 22 nodes,
- Minimize the number of OR split nodes,
- Include no more than 10 XOR, 7 AND and 4 OR decision nodes,
- Each decision nodes should have no more than 7/9 input/output sequence nodes,
- A difference higher than 15 -20 in the number of input/output sequence flows between split/join nodes is unacceptable.

Failing to adhere to these guidelines, they suggest, threatens process model conclusion validity, construct validity, internal validity and external validity. Cardosa, et al. (2006) consider research measuring complexity in business process to be a new field of research.

They suggest complexity impacts on the correctness, maintainability and understandability of process models. Their research looks for analogous metrics from the fields of cognitive science, graph theory and computer science and discusses the application of these metrics. The computer science metrics considered are the Line of Code, McCabe Cyclomatic Complexity, Halstead Complexity Information Flow Metric. Their work on cognitive science is limited to the work of Cant, Jeffrey and Henderson-Sellers (1995) who developed a conceptual model for understanding complexity in computer programming; and there research on graph theory is limited to work by Latva-Koivista (2001), Neumann (1988) and Sheppard (1990) with some rudimentary graph metrics being considered. They draw no conclusion from their research, except to state that empirical testing will be the next steps in the process.

Kluza and Nalepa, (2012) provide an overview of existing process metrics and propose a new ‘square’ metric based on the business process model and notation design (BPMN). Their approach first outlines BPMN as a set of standards against which business processes can be captured. Based on the computer sciences approach they review the work on business process metrics of Mendling, Reijers, and van der Aalst, (2010); Ligeza, (1999); Wang, et al. (2011); Grady, (1994); Monsalve et al. (2011); Khlif et al. (2010); Reijers and Vanderfeesten, (2008); Cardosa et al. (2006); Vanderfeesten et al. (2007); Conte et al. (1986); Lassen and van der Aalst, (2009); and Muketha et al. (2010). Finally they propose the use of the Durfee Square Metric (DSM) and the Perfect Square Metric (PSM). They conclude by suggesting there to be a lot of current research on business process metrics, derived from the computer industry, which has yet to be empirically validated. In this paper they propose the use of DSM as a simple but effective metrics with validation coming from future research.

Medling, Reijers and van der Aalst (2009), like most research in this domain, take information technology as the origin of process modelling and focus on the quality of business process modelling. They analyse the exiting research on the relationships between model structure, error probability and understanding, and propose seven guidelines for business model development. The basis for their analysis is: business process models have become the a focus point in the enterprise because it is the effective completion of the processes in the enterprise that enable the competitive position of the enterprise in the market place; however, effectivity is being eroded because of a lack of guidance to users on the development of effective business process models. Consequently they offer seven guidelines for the effective construction of business models.

The method they use is to build the guidelines from previous research, specifically, this research is taken from how process models are understood, the error probability of processes and the ambiguity of the process labelling. From this research they suggest seven guidelines:

- Use as few elements in the model as possible.
- Minimize the routing paths per element.
- Use one start and one end event.
- Model as structured as possible.
- Avoid OR routing elements.
- Use verb object activity labels.

(For clarity, `Model as structured as possible’ in the above list is defined as each split connector having a respective join connector).

This review of network theory in supply chain management, of alternative approaches to network theory and the earlier reviews of entropy and information theory suggest the need for greater understanding of complexity in supply chain management if the concept is to remain an effective model for understanding the structure of a multi-organisation approach to the movement of material, information and cash. Furthermore, the literature highlights the need for new approaches to methods and frameworks for understanding complexity in the supply chains. Finally, the research calls for continued experimentation and testing of methods and frameworks in what appears to be a relatively immature field of research. Given this requirement, this research will now move towards the development of an approach aimed at answering some of these calls.

2.6 Developing a revised model

To contribute to the issues raised above, a revised model is proposed for understanding complexity in the supply chain. In defining this approach, and in keeping with the internal realist ontology - an explanatory, rather than more of a mathematical approach, will be followed; developing the model from first principles with simple examples, then applying the model to more complex representations of business supply chains.

Take two dice. We know that, when throwing the dice, there is more chance of arriving at a total value of seven rather than, for instance, two. The sum or the two dice and the number of ways each can be achieved are given in Table 13:

Table 13: The number of ways the outcome of a throw of two dice can occur.

Total	2	3	4	5	6	7	8	9	10	11	12
No' of ways	1	2	3	4	5	6	5	4	3	2	1

For example, there is only one way a total of two can be achieved, by throwing two ones; but there are six ways a sum of seven can be achieved: one and six, two and five, three and four, four and three, five and two and six and one. Eleven possible results are available... two through to twelve. If the outcome of a given throw was unknown, but it was known that eleven results were possible, the result can be derived by asking binary question. Divide the eleven results into two roughly equal groups and ask if the outcome is in one of the groups. With a positive response, continue with the nominated group, a negative response means continuing with the alternative group. We saw earlier that Log to base two of the number of available outcomes defines the number of questions required, thus, using this method would require 3.459 questions:

$$Log_2 11 \ = 3.459$$

Equation 39

It is difficult to ask 0.459 of a question, further explanation is required. Imagine the group to be divided into two groups. Group 1: 2,3,4,5,6. Group 2: 7,8,9,10,11,12. In, for example, a scenario in which seven was the outcome; group two would be selected from the first binary question. The process is now repeated. Group A1: 7,8,9. Group A2: 10,11,12, and group A1 is the outcome. Another repetition gives... Group B1: 7,8. Group B2: 9 where group B1 is selected leaving a final selection between the values 7 and 8. In this case, it has taken four questions. In the scenario where 9 had been the outcome; only three questions would have been necessary; thus the value, in this scenario is between to two.

$$Log_2 11 = 3.459$$ Equation 40

Log to base two represents the minimum average number of binary questions required to identify the correct outcome which can be generalised as:

$$Qn = Log_2 N$$ Equation 41

Where Q_n is the minimum average number of binary questions and N is the number of potential states available. To follow the same format as that in the earlier review of thermodynamics, this will be labelled the macro level analysis. The readers will recall – in chapter two– the construction and correlation between macro and micro level thermodynamic entropy. A more detailed analysis of the above scenario will develop the thermodynamic equivalent of the micro (Boltzmann – Shannon) component.

In the above example, through asking somewhere between three and four questions it has been shown that a specific outcome of ‘seven’ can be found from a range of two to twelve possible solutions. What is not clear is which of the six possible solutions for the outcome seven actually occurred.

Table 14: The specific way the outcome of the throw of a die can occur.

Total	2	3	4	5	6	7	8	9	10	11	12
No’ of ways	1	2	3	4	5	6	5	4	3	2	1
Specific Way	1,1	1,2	1,3	1,4	1,5	1,6	2,6	3,6	4,6	5,6	6,6
Specific Way		2,1	2,2	2,3	2,4	2,5	3,5	4,5	5,5	6,5	
Specific Way			3,1	3,2	3,3	3,4	4,4	5,4	6,4		
Specific Way				4,1	4,2	4,3	5,3	6,3			
Specific Way					5,1	5,2	6,2				
Specific Way						6,1					

Assume, in this instance, that of the six possible ways in which the value ‘seven’ can be achieved, has been arrived at by throwing (3,4). Using the approach defined above,

$$Qs = Log_2 6 = 2.58 \qquad \text{Equation 42}$$

2.58 further questions would be required, and this will only possible if we know the number of combinations that can be brought together to equal ‘seven’. In addition, the frequency of each outcome varies. This makes the next level analysis more complex, particularly when we are looking for a simple framework that can be applied in business.

Table 15: Probabilities for the throw of two dice.

Total	2	3	4	5	6	7	8	9	10	11	12
No' of ways	1	2	3	4	5	6	5	4	3	2	1
Specific Way	1,1	1,2	1,3	1,4	1,5	1,6	2,6	3,6	4,6	5,6	6,6
Specific Way		2,1	2,2	2,3	2,4	2,5	3,5	4,5	5,5	6,5	
Specific Way			3,1	3,2	3,3	3,4	4,4	5,4	6,4		
Specific Way				4,1	4,2	4,3	5,3	6,3			
Specific Way					5,1	5,2	6,2				
Specific Way						6,1					
Probability	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36
Unavailable information	0.027	0.045	0.058	0.068	0.077	0.083	0.077	0.068	0.058	0.045	0.027

Traditionally, and as applied in previous research, using Shannon (1948), the total unavailable information (entropy) would be calculated as:

$$S = -\sum_{1}^{N} p \, log_N p \qquad \text{Equation 43}$$

$$\mathrm{S} = -\sum_{2}^{12} p \, log_{36} p \; = 0.6335 \qquad \text{Equation 44}$$

Or descriptively, if entropy is considered to be the average number of question required to ascertain the answer, the information in the system that is unavailable is 0.6335. This would suggest less than one additional question is required which appears, at face value, to be inconsistent; we see from earlier discussions that more than this number of question would be required.

Applying a simple $\log_2 f$, where f is of each of the frequency of the values 2 to 12, to the dice throwing example we arrive at the values set out in Table 16.

Table 16: The application of macro level analysis to the micro level elements.

Total	2	3	4	5	6	7	8	9	10	11	12
No' of ways	1	2	3	4	5	6	5	4	3	2	1
$Log_2 f$	0	1	1.58	2	2.32	2.58	2.32	2	1.58	1	0

The traditional Shannon method takes *log* (recall Shannon was ambiguous with regard to base, here we will use base two) of the probability of the set of occurrences; i.e in the case where seven is the value of the throw – the macro component, there is 6/36 probability of this occurrence. The $\log_2 p$ result for each of the micro components in the dice example is shown in the Table 17.

Table 17: Log base 2 for outcome probabilities.

Total	2	3	4	5	6	7	8	9	10	11	12
No' of ways	1	2	3	4	5	6	5	4	3	2	1
$Log_2 p$	-5.16	-4.16	-3.58	-3.16	-2.85	-2.58	-2.85	-3.16	-3.58	-4.16	-5.16

These values show a clear inconsistency. While it does take an absolute value of 2.58 questions to understand a systems state where the frequency is equal to six micro states and where the questions are limited to binary – yes/no – answers, the number of questions will not increase as the number of microstates decreases. But, here as the probability decreases, $\log p$ increases, which suggests that if a business question relates to the degree of say, mixed up ness, then Shannon entropy may inform; but if the question relates to understanding the state of a given system, Shannon entropy may not be particularly valuable. An alternative state rather than probability method may prove more successful.

Continuing with the two dice example and applying the traditional entropy calculation (Equation 45), the result (Table 18) – 3.27 - does not inform the understanding of the system state for four reasons: Firstly, calculating entropy to base two does not add to our understanding of state.

$$S = -\sum_{1}^{N} p\, log_2 p \;= 3.27 \qquad \text{Equation 45}$$

3.27, may inform some arbitrary measure of unavailable information, but it does not inform how we get to a greater understanding of the state of the system: 3.27 what? Secondly, if, alternatively, we calculate entropy to base 36 (Equation 46), which represents the number of available primary elements in this given dice throwing language, the result:

$$S_{u=} -\sum_{1}^{N} p\, log_{36} p \;= 0.634 \qquad \text{Equation 46}$$

is equally un-informative, 0.634 what? Thirdly, the traditional method remains lacking in units as pointed out by Tsallis (2009), and Naim (2012); thus, in either case the outcome is an arbitrary value that is relatively un-informative because of its lack of units (the need for a Boltzmann's constant) and its relativity to a what exactly it measures; it seems von Neumann may have been correct after all. Fourthly, the construct of the approach does not inform our state understanding. Specifically, in Table 18, row A equals $\log_2 f$, where f is the frequency of any given event, which represents the number of binary questions, as described above, required to determine which of the micro states is applicable. Rows B to D represent the components making up traditional Shannon entropy. None of these elements quantifiably inform understanding. As was described above, $logp$ reduces as probability increases but does not inform understanding in terms of the number of questions required to understand the state of the system. Similarly, $plogp$ does not inform understanding.

Table 18: Summary table of Shannon entropy.

	Total	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	No' of ways f	1	2	3	4	5	6	5	4	3	2	1
A	Log_2	0	1	1.58	2	2.32	2.58	2.32	2	1.58	1	0
B	p	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36
C	$logp$	-5.17	-4.17	-3.58	-3.17	-2.85	-2.58	-2.85	-3.17	-3.58	-4.17	-5.17
D	$plogp$	-0.14	-0.23	-0.30	-0.35	-0.40	-0.43	-0.40	-0.35	-0.30	-0.23	-0.14

It is worth clarifying the logic for the selection of the base of the logarithm. Hartley (1927) briefly explained, from an information theory perspective, how the selection of base two is appropriate where some form of binary decision process is in play. Following Hartley's logic it follows that the selection of the base is aligned with the number of primary elements available in the language; for instance, the English alphabet would be base twenty six. Thus, for binary questions, base two should be adopted.

Given the issues with the use of entropy, as described earlier in this chapter a revised model is proposed that takes into account the grouping of data but omits probability. We have $n \in N$ data. The data n is grouped into groups m of which there are k groupings. a is the number of n data elements in groups m such that the sum of a for all m_k groups equal n.

$$n = a_{m_1}^{n} + a_{m_2}^{n} + \ldots + a_{m_k}^{n}$$ Equation 47

For example, using the above two dice scenario, we have n=36 data elements, these are allocated to groups *m* of which there are *k* =11 groups. a is the number of n data elements (the number of ways) in groups $m_1 \ldots m_k$ such that the number of ways n = (1+ 2+3+4+5+6+5+4+3+2+1) =36.

$$\sum_{1}^{k} a_m = n$$ Equation 48

The proposed model, which will be termed 'hidden information', Q_m for the set of data n in groups *m* equals:

$$Q_m = \sum_{1}^{k} \frac{\log_2 a_m^n}{k}$$ Equation 49

Where a_m equals the frequency of the micro states that combine to form a specific macro-state. For instance, in the example above a_m equals 6 for the macro state 'seven'. Figure 22 provides a pictorial view of this explanation:

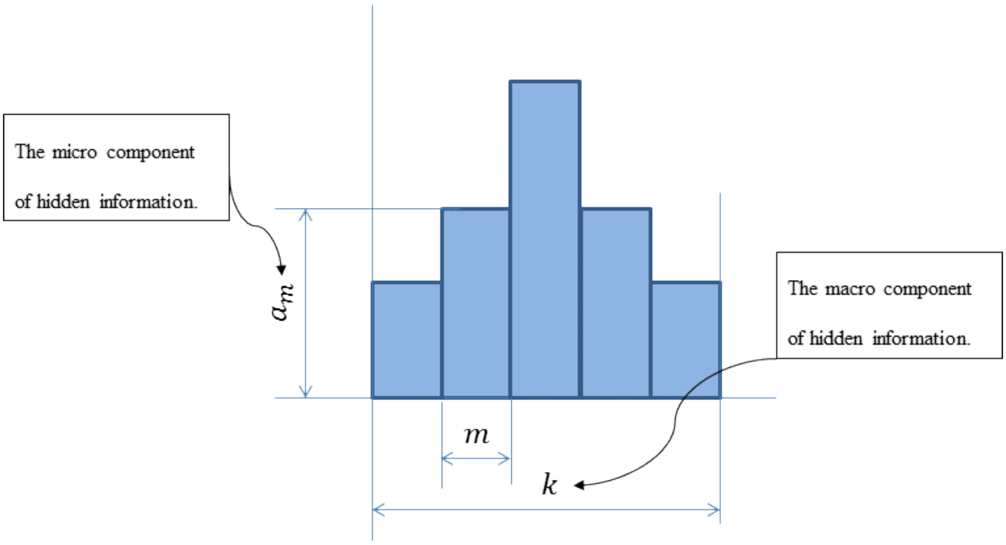


Figure 22: The micro and macro components of hidden information

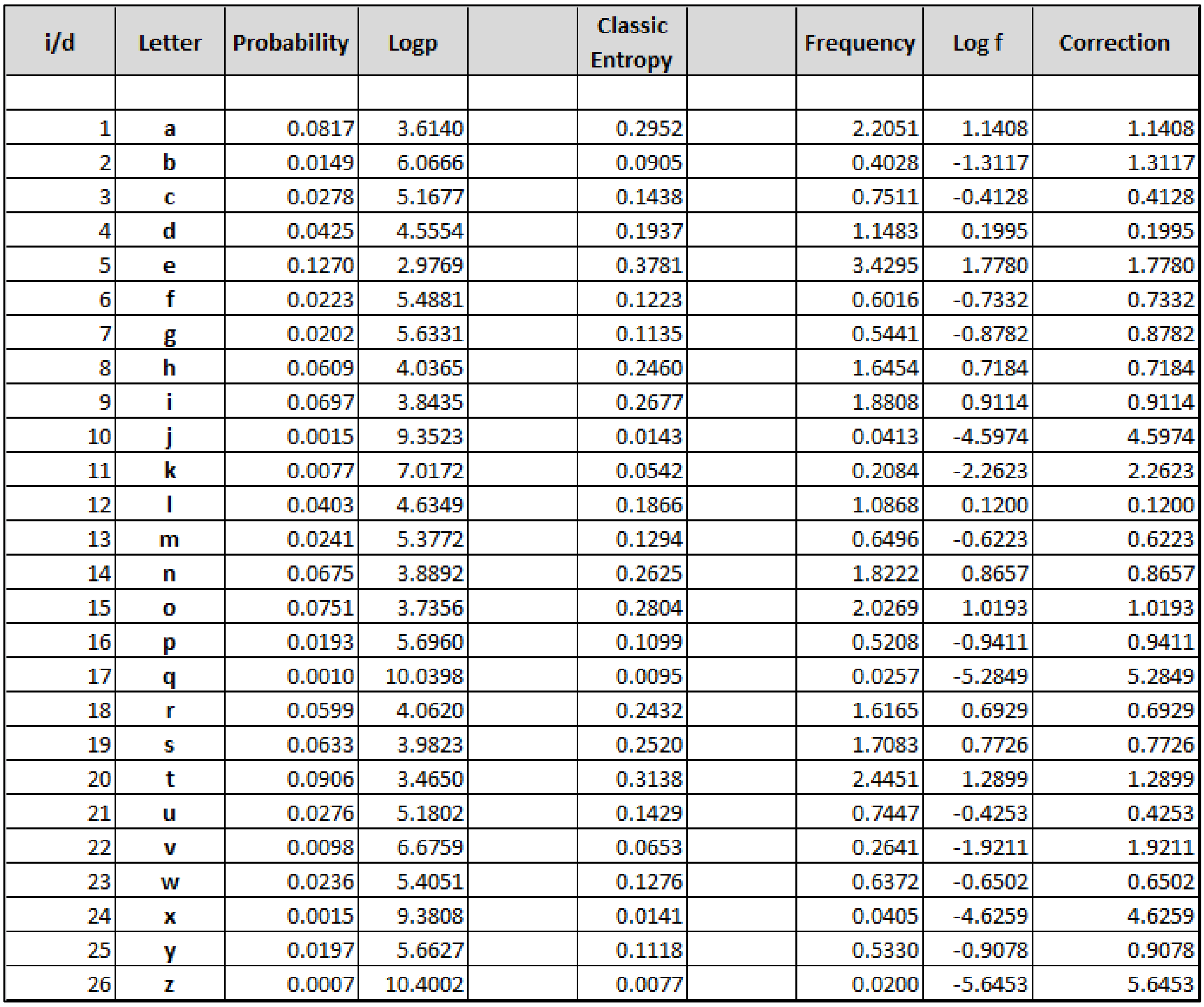
In summary, $log\ k_m$ will represent the number binary questions necessary to understand the macro state, and Q_m will represent the average number of additional questions necessary to understand the micro –state of the system. Later, these models will be used in the development of a framework for understanding supply chain complexity; however, before doing so, there is a need to move away from the field of entropy and explore these ideas from an information theory perspective.

The proposed approach follows a similar structure to the principles set out in the review of thermodynamic principles, that is, the macro (Clausius) version and micro (Boltzmann – Gibbs) version. The approach is also similar to the hierarchical construct for entropy in information theory shown in Figure 9. Like these two constructs, this approach recognises the need for understanding additional levels of granularity; however, in this case the basis is not one based on probabilities which leads to measures of mixed up ness; rather, it is based on a distinct unit of measure... that of the distance from understanding the state of a systems, measured in the number of binary questions. To explain this further, a simple worked example will be considered.

2.6.1 Worked Example

For this example the probability distribution of the letters of the English language will be used, the data has been taken from Wikipedia.

Table 19: Distribution of the letter of the english alphabet.



i/d	Letter	Probability	Logp		Classic Entropy		Frequency	Log f	Correction
1	a	0.0817	3.6140		0.2952		2.2051	1.1408	1.1408
2	b	0.0149	6.0666		0.0905		0.4028	-1.3117	1.3117
3	c	0.0278	5.1677		0.1438		0.7511	-0.4128	0.4128
4	d	0.0425	4.5554		0.1937		1.1483	0.1995	0.1995
5	e	0.1270	2.9769		0.3781		3.4295	1.7780	1.7780
6	f	0.0223	5.4881		0.1223		0.6016	-0.7332	0.7332
7	g	0.0202	5.6331		0.1135		0.5441	-0.8782	0.8782
8	h	0.0609	4.0365		0.2460		1.6454	0.7184	0.7184
9	i	0.0697	3.8435		0.2677		1.8808	0.9114	0.9114
10	j	0.0015	9.3523		0.0143		0.0413	-4.5974	4.5974
11	k	0.0077	7.0172		0.0542		0.2084	-2.2623	2.2623
12	l	0.0403	4.6349		0.1866		1.0868	0.1200	0.1200
13	m	0.0241	5.3772		0.1294		0.6496	-0.6223	0.6223
14	n	0.0675	3.8892		0.2625		1.8222	0.8657	0.8657
15	o	0.0751	3.7356		0.2804		2.0269	1.0193	1.0193
16	p	0.0193	5.6960		0.1099		0.5208	-0.9411	0.9411
17	q	0.0010	10.0398		0.0095		0.0257	-5.2849	5.2849
18	r	0.0599	4.0620		0.2432		1.6165	0.6929	0.6929
19	s	0.0633	3.9823		0.2520		1.7083	0.7726	0.7726
20	t	0.0906	3.4650		0.3138		2.4451	1.2899	1.2899
21	u	0.0276	5.1802		0.1429		0.7447	-0.4253	0.4253
22	v	0.0098	6.6759		0.0653		0.2641	-1.9211	1.9211
23	w	0.0236	5.4051		0.1276		0.6372	-0.6502	0.6502
24	x	0.0015	9.3808		0.0141		0.0405	-4.6259	4.6259
25	y	0.0197	5.6627		0.1118		0.5330	-0.9078	0.9078
26	z	0.0007	10.4002		0.0077		0.0200	-5.6453	5.6453

The validity of the distribution is not relevant or in question here and the values will be assumed to be correct. Entropy, based on the classic form equals

$$S_{u=} -\sum_{1}^{N} p\,log_2 p \ = 4.176 \qquad \text{Equation 50}$$

which provides a weighted sum of a function of probability, it is difficult to identify other measures or the information this value provides. To understand the state of the system we need to understand the complexity of the system, measured at the macro and micro level, grounded as the number of binary questions necessary to deduce a defined state. At the macro level this equates to

$$Q_n = Log_2 26 = 4.70 \qquad \text{Equation 51}$$

that is, to identify a specific letter from the list of possible letters, using the most efficient method, we would have to ask 4.7 questions. As was demonstrated above, the outcome of this macro level analysis only partly answers the question, there is a micro level component; the specific event that contributed to the frequency, what is the precise state of the system or which of the specific dice throws (1,6. 2,5, 3,4. etc) was the actual event. For the proposed measure, the average number of additional questions required is given by:

$$Q_m \ = \sum_{1}^{k} \frac{\log_2 a^n}{k} \ = 1.566 \qquad \text{Equation 52}$$

Thus the complexity, understood as the number of questions required to identify the specific state can be stated as 4.70, with an average level of non-available information specified as 1.566 questions. Or, to understand the state of the system, at the chosen level of analysis will take 4.70 questions plus and average of 1.566 additional question. The question distance is 6.266

questions. The idea of this two dimensional approach to hidden information, defined as question distance, will be the basis of the experimental approach herein.

3 Chapter Three: Research Question

Section two explained how entropy, information theory and graph or network theory have all been applied to supply chain management. It also explained the origins of entropy in thermodynamics and described the issues faced with the use of the metaphor and the transition to domains outside thermodynamics. There is increasing importance to the management of complexity in supply chain management; put simply, increases in cost, due to increases in complexity, without an appropriate increase in revenue, will be detrimental to the business; conversely, managed increases in complexity that lead to increased revenue are beneficial to the business. Hidden information, also developed in section two, may be a valuable measure of the complexity in a supply chain and a necessary component for the understanding of the diffusion of processes necessary for the successfully governed operation of the business. The research question here is: Does a process whereby hidden information is quantified provide an improved determinant of complexity of an organisations’ supply chain - that can be used to support supply chain governance and operational design – over entropy as a similar measure.

4 Chapter Four: Method

4.1 The Issues of Research

Easterby-Smith, Thorpe, and Jackson's (2012) motivation for publishing the 4th edition of their seminal book on management research was partly due to a rethinking of material on philosophy and research design. Similarly, one of Bryman and Bell's (2011) reasons for publishing a third edition of their book, Business Research Methods, was to deal with feedback on the need to embed mixed methods research as a minstream approach. Both these authors point out that an either /or approach to the more traditional positivistic or phenomenological approaches is inappropriate: 'the researcher should put effort into explaining the tendency towards one, or the other methodology'. The purpose of this section is to do just this... to construct a research context that is an appreciation of the issues that need consideration during the formulation of the research paradigm; i.e. the research environments and research perspectives in which a specific research paradigm exists. The section begins with outlines of traditional positivistic and phenomenological research methods. It then explains and positions the more complex – or grand - research theories and methods in the context of research strategy before adopting an approach for positioning these 'meta' theories. Finally the section applies a method for understanding the research approach as that of research structure and the researcher's perspectives on structure.

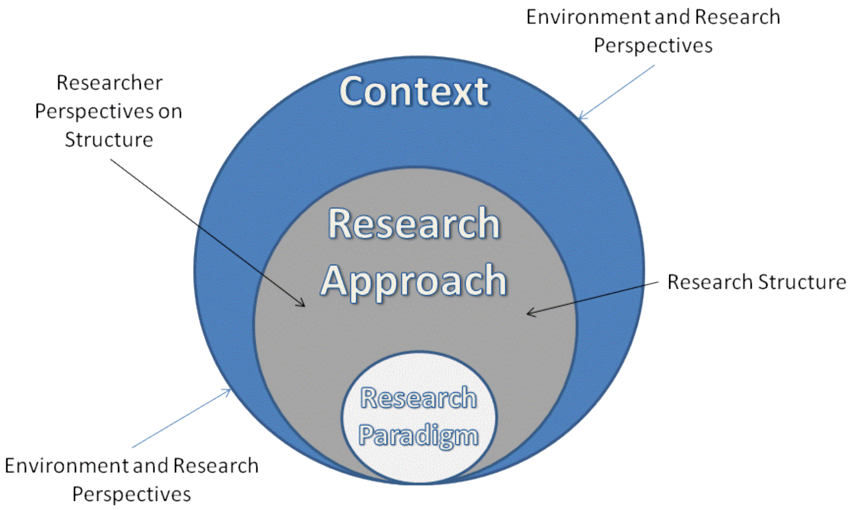


Figure 23: The Issue of Research

Classically - and generally – the epistemology in business research is classified as either positivist or phenomenological (other metaphors have been used, such as constructionist), where positivistic research is described by three general methods: Cross Sectional, Experimental and Longitudinal studies. Generally, each method can be summarised as follows. Cross sectional studies are concurrent; that is, data is taken once from comparative sources and analysed. Examples of the benefits of this method are: it is achievable in relatively short timeframes, it can have minimal subject loss and it is relatively low cost. Some issues with this type of research have been highlighted; for instance, how correlations are explained, and the influence of external variables. Experimental studies are perhaps best described by analogy: the classic laboratory experiment; where the experimental independent variables can be controlled in a systematic and procedural way. This method is usually criticised for its: lack of reality in comparison to the environment, the ability to exclude confounding variables and the role of the experimenter (Bryman & Bell, 2011; Hussey & Hussey, 1997; Barber, 1976). Longitudinal studies run – as the name suggests – orthogonal to cross sectional studies. These studies aim to analyse the dynamics of the situation to gain an understanding of the change process. Thus, this type of study is able to propose explanations of correlations. One of the criticism of the approach is that is lends itself more to phenomenological research types (Stebbins, 2006).

Similarly, phenomenological research can be described as: action research, case studies, ethnography and the feministic perspective. Action Research, which assumes the world to be constantly changing, with the research and researcher being embedded in the process, is based on a cycle of planning, acting, observing and reflecting (Lewin, 1946), and the method is closely associated with consultancy projects intended to bring change to organisations. Action research includes the following characteristics specific to the organisation: It includes a goal for the organisation. It should be recognised as a journey of learning for the organisation and be collaborative between researcher and client. This type of research must not be judged against a positivistic paradigm (Bryman & Bell, 2011). Case studies are characterised as a thorough examination of a specific unit of analysis. This could be; for instance, a company, a group, a function or process. Ethnography, based on an anthropological approach, studies the culture, symbols and rituals of a specific unit of analysis such that the social world within the scope of the research can be interpreted. The Feministic Perspective, rather than being considered as a research method in its own right, should perhaps be considered as a sub-method for other research methods requiring interviews, as it is based on a different –discursive perhaps – type of interview. This reasoning follows Bryman and Bell (2011) and Hussey and Hussey's (1995) discussions on the subject; although the latter, along with Easterby-Smith, Thorpe, and Jackson (2012) do categorise the approach as a separate method, whereas Malpas and Wake (2013) classify the perspective as a subset of critical theory (which will be discussed later). Grounded theory is characterised by the iterative nature of theory development and the combined use of inductive and deductive methods. Finally, hermeneutics, which is best seen as the opposite of the positivistic approach, focusing on understanding human action; as such, it includes an historical component intended to provide context. Grint (2000) argues the approach to be the only method by which management can operate successfully. Philosophically there is an argument put forward that hermeneutics is a separate theory to phenomenology (Smith, Flowers & Larkin. 2009) with hermeneutics being the theory of interpretation, and phenomenology the approach to the study of experience.

Other views on research have been suggested; for instance, Bryman and Bell (2011) contrast deductive and inductive research theory against an epistemological and ontological dyadic; positivism vs. interpretivism and objectivism vs. constructionism respectively. Bryman and Bell (2011) and Easterby-Smith, Thorpe, and Jackson (2012) also point out that comparison of the two research theories, as an either/or approach to research paradigm selection is inappropriate and that research is best classified as a tendency towards one or the other propositions; a view that adds further weight to the argument that time is spent reasoning a research contexts prior to the specification of a particular methodology. They highlight the work of Burrell and Morgan (1979) as being a key influencer of any given research strategy. Bryman and Bell (2011) use the work of Hassard (1991) as an example of the outcome of using the approach adopted by Burrell and Morgan (1979) which builds a multi-dimensional approach through the combination of two fundamental pairs of opposing perspectives; theories that emphasise stability and regulation versus theories that emphasise radicalism, and theories that emphasise subjectivity versus theories with an emphasis on objectivity. The multidimensional approach is developed as a quadrant:

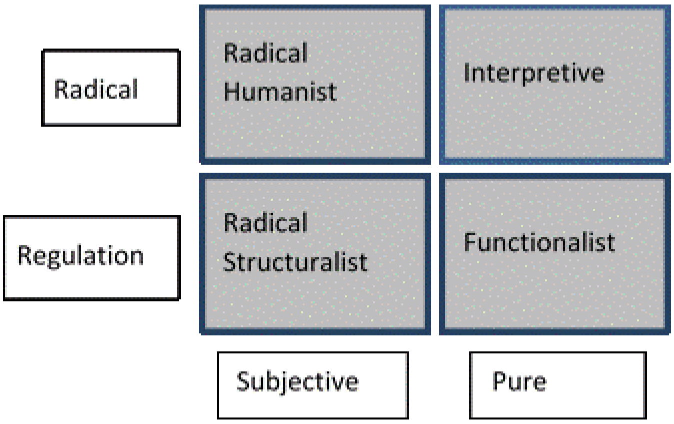


Figure 24:The multidemensional approach; developed by Burrell and Morgan (1979)

The argument proposed by Burrell and Morgan (1979) is that each research approach gives rise to different research outcomes due to the underlying assumptions. The suggestion here is that the multiple interpretations give rise to a more rounded, holistic, informative and contextual outcome. It seems reasonable to suggest that this multidimensional approach answers some of the criticisms from business with regard to the context or 'too narrow' approach adopted as a consequence of academic rigor. Epistemological diversity, within business and organisations, has -argued Buchanan and Bryman (2007) - given rise to a 'paradigm soup'.

Selecting a research paradigm is the context of the above is therefore a challenging process that needs to consider wider set – what will be labelled here as research strategy – of elements that should be considered by the researcher as those necessary for the completion of the research; thus, research strategy is not straight forward. The tendency in research is to consider the approach to research to be either positivistic or phenomenological, with the scientific world further tending towards the positivistic research type. A simple critique of the positivist approach is that it loses sight of the pluralism necessary to make the research externally valid. But this critique misses the point that research is completed in the context of the social science disciplines of sociology, psychology, anthropology and economics, and therefore pluralism itself implies complexity which needs to be understood or at least framed. Furthermore, the tendency in research is to critique from an either or perspective: positivism is criticised from a phenomenologist perspective and vice versa; or positivism is used to critique positivism using a different set of assumptions. Again, this is too simplistic. Research, from whichever perspective, adds to knowledge. Accepting that the social sciences are complex and pluralistic, a research strategy and design are needed that recognises the ethnographic context and hermeneutic interpretation within which the research resides; research therefore needs to be set in context. The next section will discuss and develop a rationale for the research design of the thesis. The section is not intended as a literature review of research methods; the purpose of the section is to develop a rationale in the context of this thesis.

4.1.1 Research Strategy Development. As already stated, the context for research strategy development is not a simple; so for clarity, the context will be broken down into three components: research environment, research perspectives and research paradigm. The researcher’s response to these components determines a research strategy. The process of reviewing the research environment, perspective and creating a research paradigm will be referred to as the research approach. The research approach is therefore the process of creating a research strategy, by responding to the environment and perspectives, and creating a research paradigm. Research environment, the research perspectives and research approach will be covered in this section; research paradigm will be covered in the following section. The first component – research environment - is represented in Figure 25:

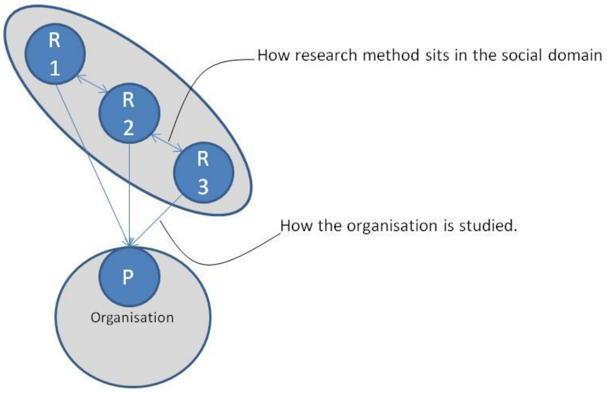


Figure 25: Research Context Component One

R1 to R3 are representative of three different research perspectives in a social environment, and P representative of a problem space in an organisation. Three key elements to the diagram exist: Firstly, how does the research perspective sit in the social environment; for instance, how does a more systemic or complex research methodology, like symbolic interactionism for instance, sit in the context of a social culture driven by the predominant social, psychological, anthropological and economic characteristics of this environment? Secondly, how is the organisation problem perceived by the researcher; e.g. is the problem based on data informed inference, an opportunity (the space shuttle disaster is a case in point) or the experience of the researcher? The author's experience in multicultural supply chain management is partly a motivation for this research. For instance, to the author, working in an Indian culture, the environment is predominantly one where validity comes from data analysis in case studies and the cross sectional study of similar cases which generate the business case for change management, but with limited application of appropriate research methodologies – lots of logical 'converse accidents', so to speak. Comparatively, the author's experience working in the Arabic culture was one where the dominant environment was one based on action and ethnographic ontologies; hence – partly - some of the motivation for this research. Finally, how is the organisation to be studied in the context of the first and second elements, positivistically, phenomenologically or through a different approach driven by the route through the research environment and research perspective?

The second component of research context is the research perspective; the set of baseline assumptions made by the researcher concerning the purpose of the research and the research epistemology. For instance; the purpose of the research may be to research problems identified by organisation management (Gummesson, 2000); alternatively, the purpose of the research may be only to add to knowledge and not be influenced by the specific needs of an organisation (Burrell, 1997). Also, the researcher's assumptions may determine that knowledge will only be gained through evidence (Rosseau, 2006; cited in Bryman and Bell, 2011), or that knowledge will be gained through a pluralistic approach (Learmonth, 2008). It is this pluralistic approach that has given rise to what Bryman and Bell term grand theories, examples being: Symbolic Interactionism, Critical Theory, Structural Functionalism and Abstraction. Easterby-Smith, Thorpe and Jackson (2012) offer Critical Theory, Critical Realism, Feminism, Hermeneutics, Postmodernism, Pragmatism and Structuration Theory as 'other' research philosophies. To be frank, the extent to which grand theories are described and contextualised in Bryman and Bell (2011), and Easterby-Smith et al. (2012) does not do them justice. Malpas and Wake (2013) offer a much more detailed view on critical theory positions and their composition in structuralism and semiotics, narratology, marxism, poststructuralism, historicism, psychoanalytic criticism, deconstruction, feminism and others. That said, Malpas and Wake offer details on content and less on the application to business research. Bhaskar (2008) offers a detailed, complex and insightful view of critical realism where he challenges the traditional constructs of research reliability and validity by disconnecting causal laws from methodology in nature by asserting the researchers transcendental perspective of intransitive causal structures and generative mechanisms in nature. In doing so he argues the creation of the closed system necessary for research validity to be only valid from the boundary conditions of the researchers mechanism of creating the transcendental perspective of the generative mechanism thus, only in the given research case, creating a naturally false correlation between generative mechanism and causal law. Bhaskar's (2008) view on critical realism is important to this research for three reasons. Firstly, Bhaskar's critique of the more traditional approaches to research adds weight to the need to explain and justify the research strategy and approach used herein. Secondly, the subject of this research is – as will be seen later – aligned with the concept of information or business architectures being subject to aggregation or consolidation that is descriptive of the intransitive causal structure and generative mechanisms discussed by Bhaskar (2008). Thirdly, it would be somewhat hypocritical to acknowledge the critique of reliability and validity and recognise the concept explained by Bhasker (2008) in the content of the research, and then ignore his critique and approach in the methodology. Consequently, inputs from critical theory and critical realism will be taken throughout this research. The importance of the approach presented by Bhaksar (2008) can be seen in Morgan (2006) and Easterby-Smith, Thorpe and Jackson (2012). Morgan (2006) describes organisation using eight key metaphors:

- Machines
- Organisms
- Brains
- Cultures
- Political Systems
- Psychic Prisons
- Flux and Transformation
- Instruments of Domination

The important point, he explains, is the description of an organisation using the structured set of metaphors set out in his text is invalid and, like the point made by Easterby-Smith et al. (2012) on business research, that organisations adopt all of these descriptions concurrently. It is an ‘and’, not an ‘or’, to a greater or lesser extent. The point made is the same: Organisations adopt multiple perspectives concurrently; it is this concurrency that provides the rich divergence, innovation and creativity in organisations. Business research requires clarity, rather than selection to the exclusion of others, of the many research ontologies and epistemologies, and in this thesis, the internal realist perspective aims to provide a richer content than a pure realist/positivist perspective.

That said, the literature offers three views on the incorporation of grand theories: Firstly Merton (1967) suggests grand theories to be too abstract or general to guide the researcher in deriving empirical data such that the researcher could make use of the data relative to the theory. Secondly, there are research examples where grand theories have been used successfully in research: Giddens (1984, cited in Bryman and Bell, 2011) is a case where the use of a grand theory - structuration theory to be specific - has been useful in research terms. Thirdly, referring to some of the same grand theories but under the heading problem structuring methods, Mingers and Rosenhead, (2004) characterise these methods as necessary for the understanding of problems where multiple actors, multiple perspectives, conflicting interests, important intangibles and key uncertainties exist. Referring to Ackoff (1979) and Checkland (1985) they propose these grand or problem structuring methods to be ‘strategic’ by setting assumptions and context. Using Ferris, (2009) this thesis postulates these three views can be reconciled by relating the theories to the nature of knowledge required. Ferris, summarising the work of Ryle (1948), Biggs (1999) and Nissen (2006), proposed knowledge to have three forms, declarative, functional and procedural. The three forms have two clear distinctions: Declarative knowledge is related to elements that can be written and taken as a given; for instance, a school multiplication or log table. Functional and procedural knowledge can be characterised as understanding the ability to understand and, if appropriate, act on that understanding; a car mechanic or a medical doctor might be a good example. Grand theories, from the critiques and positioning above, can be said to be theories aimed generally to add knowledge as understanding whereas mid-range theories can be generally related to adding declarative knowledge. Using this assumption the purpose of research can then be defined as intending to contribute to declarative knowledge and understanding knowledge by using the most appropriate combination of grand and mid-range theories.

The above outlines the high level research context and intent, but it lacks process; a route to arriving at a defined strategy for a research problem; that is, there is a need for a research approach.

4.1.2 Research Approach. A research approach can be viewed as having two components; the approach to the structure of the research, which can be seen as what the research is intended to do; and the perspective of the approach to the research intent from the researchers viewpoint. Thus, researcher perspective covers both the content perspective and the approach perspective – this is the ‘double hermeneutic’ issue referred to in the literature on phenomenological research, see, for instance, Smith et al. (2009). Research structure and researchers perspective on structure are discussed below.

4.1.3 Research Structure. Ferris (2009) recognised the need for a research strategy. He developed an approach based on a tabulated question set originally developed by Varro (it is worth explaining Varro’s work is referred to by Augustine in *De Civitate Dei* XIX:1), Varro’s original, and Ferris’s proposed taxonomy are set out in Table 20.

Table 20: Varro's (left) and Ferris's (right) taxonomy of research characteristics.

Dimension	Possible Categories
Desiderata	Pleasure
	Repose
	Combination of pleasure and repose
	Primary natural blessings
Relation to Virtue	Virtue is to be desired as the goal
	Desiderata to be considered to provide virtue
	Both virtue and desiderata to be desired
Beneficiary	Self
	Others
View of Certainty of Knowledge	Old academy – views have certainty
	New academy – views are uncertain
View of Tradition	Hold to tradition of philosophy
	Cynics
Objective of Life	Look for life of leisure
	Look for life of business
	Look for life of leisure and business

Dimension	Ref	Possible Categories
Desiderata	D1	Develop the theory of the field
	D2	Develop the practice of the field
	D3	Develop the theory and practice of the field
Relation to Knowledge	K1	Knowledge is the desired goal
	K2	Knowledge is desired for practical application
	K3	Both knowledge and application are desired
Person who benefits	P1	Researcher
	P2	Others
View of Certainty of Knowledge	C1	Knowledge is certain and absolute
	C2	Knowledge is relative and contingent
View of Tradition	T1	Conform to general pattern of discipline
	T2	Challenge or rejection of tradition
Objective of Life	O1	To enjoy knowing – Life of leisure
	O2	To enjoy practice – Life of Business
	O3	To enjoy both knowing and practice

It is worth noting that Varro's taxonomy allows for 288 philosophical categorisations, Ferris's 216; highlighting how the selection of research strategy and consequent methodology need to be thoroughly considered.

Ferris's work then builds towards a research strategy by providing a taxonomy of questions for characterising the structure of the research. This taxonomy provides an approach to reviewing the research intent against a predetermined question set in order to inform an appropriate research paradigm – the third component of the research strategy. To arrive at an appropriate paradigm the scope of this research has been considered against this taxonomy, the results of which are set out in Table 21.

Table 21: Taxonomy for the selection of a research method. Adapted from Ferris 2009.

Dimension	Question	Answer specific to this thesis
Desiderata	Is the proposed project intended to make a significant contribution to the theory of the field?	This thesis is intended to make a significant contribution to the theory of the field.
	Is the proposed project intended to make a significant contribution to the practice of the field?	This thesis is intended to contribute to the practice of the field.
Relation to Knowledge	Is the knowledge expected in the project primarily desired for its intrinsic value?	The knowledge expected is desired for its intrinsic value and it's instrumental value

Dimension	Question	Answer specific to this thesis
	Is the knowledge expected in the project primarily desired for its instrumental value in order to achieve something else?	The knowledge expected is desired for its intrinsic value and it’s instrumental value
Person who benefits	Is the primary beneficiary of knowledge expected in the project the researcher?	The primary beneficiary of the research is the researcher to a lesser extent
	Is the primary beneficiary of knowledge expected in the project people other than the researcher?	The primary beneficiary of the research is other researchers and practitioners
View of Certainty of Knowledge	Does the proposed project presuppose that the knowledge to be developed concerns matters which objectively exist?	The project presupposes that knowledge concerns matters which objectively exist within a context bounded by the individual or group.
	Does the proposed project presuppose that the knowledge to be developed concerns matters which are constructs of the community?	The project presupposes that knowledge concerns matters which objectively exist within a context bounded by the individual or group.

Dimension	Question	Answer specific to this thesis
View of tradition	Does the proposed project presuppose that the existing framework of the field should be used as a foundation?	The project presupposes that an existing framework of the field be a foundation.
	Does the proposed project presuppose that the existing framework of the field should be rejected and vigorously challenged?	The project presupposes that an existing framework of the field be a foundation on which new frameworks can be built

Given the answers set out in Table 21, the next section will review the second element of the research approach, that of the researcher's perspective on structure.

4.1.3.1 Researchers Perspectives on Structure. Earlier it was suggested that the researcher does not act independently of the research environment, the reader may recall Bhaskar's (2008) view that the researcher is in fact the creator of the environment. It was also suggested that, from the perspective of the author, the motivations for this thesis are supported by anecdotal evidence of supply chains demonstrating complex and systemic characteristics. Jackson and Keys (1984) point out that while there will be general consensus on the principle that problems in simple systems are easier to solve, it is the researcher and the research strategy that determine if the problem is simple or complex. They refer to Vemuri (1978) for three reasons for considering problems to be complex, these are reproduced below: 'Firstly, in complex –etc. systems, not all of the attributes of the elements of the system will be directly observable. As a result it is difficult to understand the nature of the systems completely. The causes of the problem may be hidden, and this will impede the ability of the problem solver to identify solutions. Secondly, in complex – etc. systems, even if laws can be established relating the actions of different parts of the system, they will invariably be only probabilistic in nature. Thirdly, complex –etc. systems evolve over time. This evolution stems, in large part, from the fact that such systems are in constant interaction with their environment'.

Mingers and Rosenhead (2004) refer to the work of Jackson and Keys (while setting out six key characteristics of complex problems:

- The problem is less structured.
- There are multiple actors.
- There are multiple perspectives.
- There are conflicting interests.
- Important intangibles exist.
- Key uncertainties exist.

Edwards and Yearworth (2011) referred to Jackson and Keys (1984) and Mingers and Rosenhead (2004), in their explanation of the need for complex research methods for systems engineering doctoral students. Specifically they highlight soft systems methodologies and in doing so they raise four 'axioms' representing the characteristics of an environment where soft systems methods would be applicable:

- Where problems do not exist independent of human beings.
- Where the problem space has potential sub-problems that can interact.
- Where solutions are 'intellectual constructs' and not isolated from the rest of the system.
- Where benefits are most likely to come through sharing of perception, persuasion and debate.

They also suggest, in these situations, researchers cannot sit as ‘objective outsiders’ of the system, as is the case for harder systems. In summary, for a system to be considered to be complex, it should demonstrate three key characteristics: Firstly, problem structuring is made complicated by problem integration; i.e. it is the integrated meta-level problem that is greater than the sum of its sub-problem components. Secondly, problem perceptions can only be aligned through human interactions. Thirdly, as a corollary to the previous two points, iteration through the process of problem resolution is likely; that is, the act of analysing the problem and defining a potential solution is likely to change the problem. For supply chain management to be considered a complex problem, from the point of view of the researcher and as stated by Johnson and Keys (1984), there must be some alignment between these characteristics and those manifest in the supply chain management topic.

A definition of, and search for, a set of characteristics that align the nature of complexity with the complexity found in supply chain management is not the purpose of this research. However, the thirty years’ experience of the author and the consequent motivation for this research – the systemic perspective – recognise significant anecdotal evidence, from the engineer to order, make to order industries, that supply chains do demonstrate complex characteristics as defined above. Based on the author’s standing assumption that most statements ‘define what they are not’, the tendency of the author, in terms of problem structuring and solving is to take a multi strategy perspective that can is shown in Figure 26:

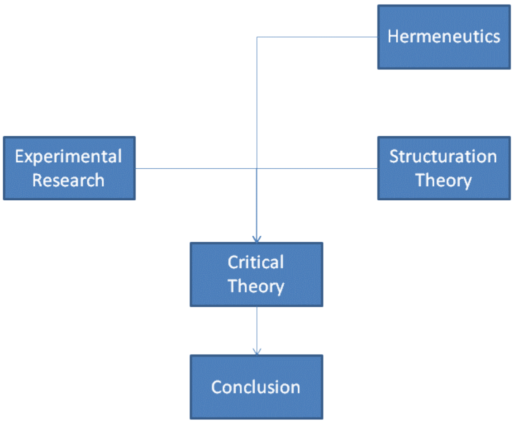


Figure 26: The Researchers Perspective on Problem Scoping and Solving.

The term 'defines what they are not' is worthy of a small digression to explain the meaning. Earlier, when reviewing the work of Hartley (1927) it was be explained how, for instance, if a sentence makes reference to a dog, the sentence then excludes all other animals; i.e. all animals that are not dogs are excluded. This exclusion principle is what is meant in the term above (Readers interested in this approach are referred to Crisp and Turners (2010) discussion on social categorisation).

Earlier in this section the general issues of research were considered. The issue of research strategy was discussed, highlighting the issue of problem definition in the environment of grand or problem structuring theories and the better known theories applicable to the 'mid-range', classical approaches to research. The motivation for this research has its basis in the assumption that supply chains are a complex system; as such, some of the grand theories appear to have some applicability to the research. However, as stated in the critiques of grand theories, using these theories would make it difficult to define a specific problem statement. To resolve this conundrum, without losing sight of the wider perspective, the research problems defined in this research will be considered to be 'problem dimensions' of a wider – systemic - problem set. As such, we have to accept that knowledge generated from this research will be functional; i.e. the knowledge will add to understanding rather than be declarative knowledge.

4.1.4 Summary

The intention of the sections on research issues, research strategy and research approach was to outline the issues associated with research methods and explain how, by structuring an approach to understanding the research environment and research perspectives, and to the formation of a research strategy and approach, a context could be provided to a specific research paradigm.

To summarise:

- Research is not as simple as deciding over a positivistic or phenomenological approach.
- There are methods and theories that recognise a more complex approach to research design.
- The complex methods can be used to frame a specific research problem.
- Applying a structured approach to framing the research problem informs the selection of an appropriate paradigm.
- The researcher’s perspective on both the research problem and research ontologies and epistemologies is an equally key determinant of the research context and therefore a key input to the research paradigm.

With this information, the next section will develop a research paradigm specific to this thesis.

4.2 Research Paradigms

The purpose of this section is to frame a research paradigm specific to this thesis. Framing the paradigm will consist of three sub-sections. Firstly Ferris's 2009 taxonomy will be used to frame the intent of the thesis. An explanation of research purpose and research process will complete the context and an explanation of the specific paradigm will then be provided.

4.2.1 The Context of this Research. Earlier Ferris's taxonomy was explained. The taxonomy specific to this research is reproduced in Table 22. The focus in this section is the answer to each of the questions which describes the research intent and presuppositions, and serves to inform the selection of an appropriate research process, and frame the research paradigm.

Table 22: Ferris's taxonomy reproduced from table 21.

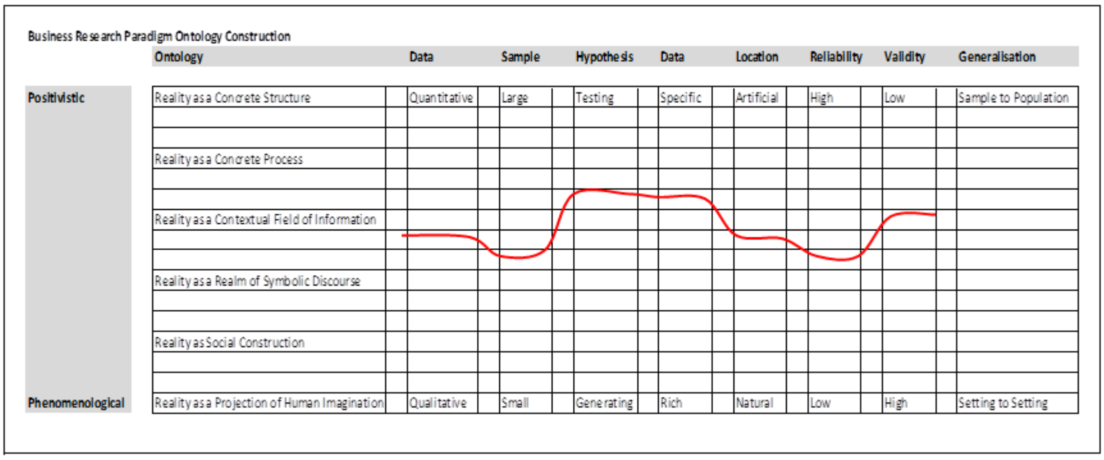
Dimension	Question	Answer specific to this thesis
Desiderata	Is the proposed project intended to make a significant contribution to the theory of the field?	This thesis is intended to make a significant contribution to the theory of the field.
	Is the proposed project intended to make a significant contribution to the practice of the field?	This thesis is intended to contribute to the practice of the field.
Relation to Knowledge	Is the knowledge expected in the project primarily desired for its intrinsic value?	The knowledge expected is desired for its intrinsic value and it's instrumental value
	Is the knowledge expected in the project primarily desired for its instrumental value in order to achieve something else?	The knowledge expected is desired for its intrinsic value and it's instrumental value

Dimension	Question	Answer specific to this thesis
Person who benefits	Is the primary beneficiary of knowledge expected in the project the researcher?	The primary beneficiary of the research is the researcher to a lesser extent
	Is the primary beneficiary of knowledge expected in the project people other than the researcher?	The primary beneficiary of the research is other researchers and practitioners
View of Certainty of Knowledge	Does the proposed project presuppose that the knowledge to be developed concerns matters which objectively exist?	The project presupposes that knowledge concerns matters which objectively exist within a context bounded by the individual or group.
	Does the proposed project presuppose that the knowledge to be developed concerns matters which are constructs of the community?	The project presupposes that knowledge concerns matters which objectively exist within a context bounded by the individual or group.
View of tradition	Does the proposed project presuppose that the existing framework of the field should be used as a foundation?	The project presupposes that an existing framework of the field be a foundation.
	Does the proposed project presuppose that the existing framework of the field should be rejected and vigorously challenged?	The project presupposes that an existing framework of the field be a foundation on which new frameworks can be built

4.2.2 The Purpose of this Research. The purpose of research can be classified as exploratory, descriptive, analytical and applied or basic. The intent of exploratory research is to provide insight into a subject such that further research and insight can be developed. The outcome of exploratory research is unlikely to be conclusive; rather, it is likely to assess and investigate patterns and hypotheses; rarely providing conclusive answers. Descriptive research is intended to describe phenomena as it exists. The approach goes further than exploratory research in that it would normally attempt to describe the characteristics of the research issues. Analytical research moves research maturity further still with the development of causal attributes being developed for the identified issues. Thus, analytical research aims to understand the variables that impact on the identified issues. Predictive research goes even further still by predicting future outcomes for given variables. Hussey and Hussey (1997) suggest predictive research covers the ‘how, why and where’ answers.

To meet the intent of this research, ideally, the purpose of the research would be analytical. The author recognises the limitations that may occur with the availability of relevant data. In these cases, a need for further research will be highlighted.

4.2.3 The Process of Research. It was highlighted earlier that research paradigms are not simply a selection between positivistic or phenomenological research types. It was also highlighted that any selection between the two is not a selection between one and the other, rather it is a selection on the degree to which one or the other is selected. The degree to which one or the other is selected is shown in Figure 27 as a line through a matrix construct developed from Hussey and Hussey (1996).



Business Research Paradigm Ontology Construction

	Ontology	Data	Sample	Hypothesis	Data	Location	Reliability	Validity	Generalisation
Positivistic	Reality as a Concrete Structure	Quantitative	Large	Testing	Specific	Artificial	High	Low	Sample to Population
	Reality as a Concrete Process								
	Reality as a Contextual Field of Information								
	Reality as a Realm of Symbolic Discourse								
	Reality as Social Construction								
Phenomenological	Reality as a Projection of Human Imagination	Qualitative	Small	Generating	Rich	Natural	Low	High	Setting to Setting

Figure 27: Research ontology, developed from Hussey and Hussey (1996).

4.2.4 Research Paradigm.

To begin to build a specific research paradigm a summary of the author's position on research strategy and approach is as follows.

The positivist or realist perspective creates information by nature of the natural sciences phenomena; realism is validated by statistical data and criticized for omitting reliability or context. Conversely, the nominalist/constructionist perspective creates information through interpretive analysis of rich data to infer generalization. Grand theories bring a different perspectives; hermeneutics, critical theory and critical realism in particular, and make the point that knowledge is:

- Formed internally.
- Double hermeneutic (the author forms a meme of knowledge, writes it down, and the reader interprets that knowledge and assigns attributions to the knowledge from his or her perspective).

Interpretive phenomenological analysis (Smith et al., 2009) goes further by investigating 'how' individuals interpret. The social sciences explain the issue of attribution, and how errors occur. Logically, attributions are assigned and some of these assignments are done in error; thus social psychology's set of defined attribution errors. Hypothetically there is no 'normal' assignment, attribution assignments are not completed against a standard model, they are assigned against an individuals' perspective. As a consequence, attribution error can only be described as an assignment that is not within boundary conditions considered to be normal, and all attributions are to some degree, erroneous.

To an extent Aristotelian or Boulean logic provides a set of rules that can be applied to the assignment of attributes. Data, news, opinions, facts and other knowledge all class as information which is bounded by the amount and interest assigned to the information by the recipient. As a consequence, what an individual knows can be suggested to be the attributions assigned to all the information selected by the individual in which he or she has sufficient interest in which to care to assign an attribute. This argument is similar to those proposed by, for instance Husserl (circa 1927) and Heidegger (1962) who, according to Smith et al. (2009) were the main protagonists of interpretive phenomenology from an epistemological and ontological viewpoint respectively.

There is a clear argument for the principle that knowledge is constructed from the attributions assigned by the individual to only the information they have available. This is an important issue for any research, and from the perspective of this research it is important for three main reasons: Firstly, there is an internal realist perspective to the external validity of the research question. Secondly, the methodology attempts to retain an internal realist – rather than realist – perspective throughout the text. Thirdly, how the findings of the research are considered is approached from this perspective rather than the realist, relativist or nominalist perspectives. Consequently, the epistemology tends to follow a weak positivistic position and the text itself takes time to bound each point made to inform an internal realist perspective.

Easterby-Smith et al. (2012) tabulate research types for give research assumptions. Their table is replicated in Table 23.

Table 23: Methodological implications of different Epistemologies. Reproduced from Easterby-Smith, Thorpe and Jackson (2012).

Ontologies	Realism	Internal Realism	Relativism	Nominalism
Epistemology Methodology	Strong Positivism	Positivism	Constructionism	Strong Constructionism
Aims	Discovery	Exposure	Convergence	Invention
Starting Point	Hypothesis	Proposition	Questions	Critique
Designs	Experimental	Large Surveys: Multi-cases	Cases and Surveys	Engagements and Reflection
Data Types	Numbers and Facts	Numbers and Words	Words and Numbers	Discourse and Experience
Analysis/Interpretation	Verification/ Falsification	Correlation and Regression	Triangulation and Comparison	Sense Making: Understanding
Outcomes	Confirmation of Theories	Theory testing and Generation	Theory Generation	New Insight and Actions

Easterby-Smith, Thorpe and Jackson (2012) use different descriptions for the ontological and epistemological dyadic. Ontology is referred to as realism and nominalism; epistemology is referred to strong positivism and strong constructivism. For clarity, while recognising the value in Table 23, this thesis will continue with the objectivism – constructivism and positivistic – phenomenological labelling for the terms.

The approach herein to contextualising the issue of research, while seemingly laborious, is not unique. The theoretical physicist Lee Smolin (2013) argued a similar perspective for the content of his book ‘Time Reborn’. The background on the issues of research, the adoption of a research strategy through the application of a research approach allows the paradigm for this research to be clearly articulated. To be specific, and with reference to Table 23, in the context of a wider research strategy, this thesis will adopt an objective/positivistic approach with the design of the research being experimental. The experimental design will be in the form of a laboratory experiment, rather than a field type experiment. The reason for this selection is that there is a high risk to construct and predictive validity, and to the reliability of the research if field data variables are included at this stage. Further research will be required to test the outcome of this research against field data or alternative research paradigms. That said, the experience of the author in a range of industries is that more time and effort is spent in organisations dealing with issues brought about because of the internal realism construct described above. As a consequence this research will take an ontological position that can be best categorised as internal realism. However, as explained in chapter four the wider context will be provided through a thread of critical realism which will allow for a phenomenological or natural thread to run through the research without specific observations being available. In this context a research methodology will also be set out in chapter four.

The rest of this thesis is set out as follows. The research methodology will now be explained. Chapter five will construct and complete the analysis and chapter six will discuss the major findings, meaning and importance, relation to similar findings, alternative explanations, relevance, limitations and further research. Chapter seven will draw conclusions from the research.

4.3 Research Methodology

The Supply Chain Operating Reference Model^{TM} is a well-established business process model used in the management of supply chain processes. The framework consists of five main subsystems: Plan, Source, Make, Deliver and Return. Each sub-system consists of a differing number of processes which are categorised into two groups, operational and enabling. Each process consists of a set of connected activities and a set of links to other processes (The full structure of the supply chain operating reference model is provided in appendix C). This framework will be used as the basis for the process model in this research.

Using this framework a network adjacency matrix will be constructed with the business processes represented as nodes on the horizontal and vertical axes. The business process connections being represented as links by adding the value one at each relevant activity intersect. Following network theory, in the scenario where the process connection is bidirectional this will be represented with the value one being added to the intersects in both directions (i, j) and(j, i). With the processes represented on the horizontal and vertical axes, the bi-directional link will be represented as the value one at intersect A,B and at intersect B,A, as shown diagrammatically in Figure 28.

	A	B
A		1
B	1	

Figure 28: Bi-directional process connector for processes A and B.

In the scenario where a business process link is circular (self-edging in graph theory terminology); i.e. where the link goes back to the same process, this will be represented as a value one at the intersect of the same process on the adjacency matrix diagonal; for instance, if process A had a circular link, this would be represented as a value at the intersect A,A. In the scenario where a link to a process is either to or from a process that is external to the SCOR process model, an additional intersect will be added representing the link between the SCOR process and the external process, this will be done for incoming and outgoing links.

From the constructed adjacency matrix, measures of process in degree and out degree will be derived for each process and process grouping. For clarity, the in degree of a process is the number of process links coming in to the given process, this will be represented as the horizontal sum for each row in the matrix. The out degree is the number of links out of the process, this will be represented as the vertical sum for each column in the matrix.

It is important to the construct validity of this research, to address why this methodology has been selected, and specifically why the in degree/out degree has been selected as an independent variable. Business process models are not new phenomena. As a serial set of connected activities, they are well established in industry; forming part of a normal functional governance framework. Increases in the number of serial processes of a business add to the governance scope within the boundary of a function, as such the process remains constrained within the existing functional construct. Demands on businesses are, however, changing, there is a greater need for process flexibility, there is a greater demand for total solutions rather than a product only solution; there is a greater demand for supply chains to align with specific market segments and, at the same time, respond to demands in an integrated and agile model. Finally, there are trends of globalisation as industry looks to find a way to lower costs and satisfy greater demand from growing economies, all of which demands a more defined inter region and inter organisational construct. Consequently it is not so much the serial processes that measure the operating complexity of the supply chain; more is it the interplay between the processes that adds to the flexibility in the business, the agility of the business and the integration of a global business. As such it is the inter-functional interplay – the links between different activities and processes – that is adding to the complexity of a business; hence it is this phenomenon that is being considered as the independent variable against which the hypothesis will be tested.

The adjacency matrix will be constructed for the processes in the SCOR model (185 processes) plus an additional category for the links connecting into or outside of the SCOR model. For each of the processes the, number of outgoing links will be used as a measure of the interconnectedness of the process. In line with graph theory, these links will be referred to as the 'out degree' of the process and will be represented as k_{out}. For each process, entropy and hidden information will be calculated. Entropy and hidden information will also be calculated for levels zero, one, two and three levels of aggregation. Details of the exact aggregation construct are explained and tabulated in the early part of the next chapter, chapter 5.

The validity of the measure of hidden information, in the laboratory experiment, needs to be tested against scenarios designed to be representative of actual business models. Two business scenarios will be validated through field research. The field research data will be collected through structured interviews. The structured interview approach will follow the process and question set laid out in appendix D (the question set was also made available through a subscription based questionnaire at www.surveymonkey.com/s/3HNGFNH). The structured interview process described in appendix D has been selected for several reasons: Firstly, as described in 0, presentation of the information required opens up the information to the social domain, the multiple perspectives of the multiple actors (Mingers & Rosenhead. 2004), and the challenge of double hermeneutics (Smith, et al., 2009). Presenting the information requirement in, for instance, a simple questionnaire would open the risk that the information provided is significantly influenced by the contributing actors' perspectives; for instance, in the past, they may have had some bad experience in using this particular model and, as a consequence, be reluctant to, or not put much thought into constructing the answers. Secondly, the multiple actors and perspectives, would likely lead to attribution error in one form or another, risking the validity of the feedback. Thirdly, not acknowledging the fact and consequential risk that attributions may not be directly observable contradicts Vermuri (1978) and the principles of an internal realist perspective to research. Fourthly, as pointed out by Smith et al. (2009) 'because of exposure to market research and popular questionnaires' people need more engagement in the process in order to extract fuller information.

In keeping with the interpretive input and to provide input based on field experience, a small number of interviewees were selected. Their selection, by the author, was based on their knowledge of supply chain management, for instance, their knowledge of supply chain management operations, supply chain management organisation structure design and their breadth of understanding on supply chain management design options and the use of third party logistics providers. To provide differing perspectives on the content of the two scenarios, the interviewees where chosen from major businesses in different industry sectors: Aerospace and Defence, Pharmaceuticals, Life Sciences and Consulting.

A method for testing the reliability of the agreement between the field research contributors is required. The method needs to allow for multiple contributions and binary selections; Fleiss kappa (F_k) will therefore be used as the determinant of the reliability of agreement among the contributors where:

$$F_k = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}$$

Equation 53

where

$$\bar{P} = \frac{1}{N}\sum_{1}^{N} P_i \qquad \text{Equation 54}$$

and

$$\bar{P_e} = \sum_{1}^{k} p_j^2 \qquad \text{Equation 55}$$

with N equalling the number of questions and k the number of assignable categories.

Data from the structured interviews will be consolidated and aggregated to represent two business scenarios. The two scenarios will include the processes defined in the aggregated field research data. The data will be categorised under a prime contractor model and a subcontractor model. In the prime contractor model a business model is built where all the supply chain processes are utilised in the business; in addition, a second set of processes are added to the business model to represent the additional activity necessary for the prime contractor model. In the subcontractor model the assumption is that the business operates as a completer of activities that are planned and sourced by a prime contractor; hence, in this model, the element of the supply chain model will represent a reduced set of business processes. For each scenario the out degree will be used to determine entropy and hidden information at the four levels of aggregation. k_{out} is therefore the independent variable for this laboratory experiment. Appendix B explains in more detail how the supply chain operating reference model is constructed and how this construct has been used to create the adjacency matrix. Appendix C details the complete tabulated supply chain operating reference model.

5 Chapter Five: Analysis

5.1 Construction of the adjacency matrix and the business scenarios

5.1.1 Basic Adjacency Construct

Chapter four explained how the SCOR^{TM} model was to be analysed as an adjacency matrix, using K_{out} as the independent variable across the hierarchical aggregation structure of the model. The SCOR^{TM} model assigns processes to structural levels one to four, with level four usually being specific to any given organisation. For the purposes of this research levels one to three are used with an additional level (zero) being included to represent the results at the top level (i.e. at the SCOR^{TM} level). The SCOR^{TM} model has therefore been constructed in the form of an adjacency matrix with the horizontal and vertical axes representing the defined processes, grouped as defined in the model, into a hierarchical structure. The axes structure is defined in the Tables 24 - 28.

Table 24: Axes structure for SCOR plan.

Level 0	Level 1	Level 2	Level 3
Plan	Plan (Operations)	SP1	SP1.1, SP1.2, SP1.3, SP1.4
		SP2	SP2.1, SP2.2, SP2.3, SP2.4
		SP3	SP3.1, SP3.2, SP3.3, SP3.4
		SP4	SP4.1, SP4.2, SP4.3, SP4.4
		SP5	SP5.1, SP5.2, SP5.3, SP5.4
	Plan (Manage)	SEP	SEP1 through to SEP10

Table 25: Axes structure for SCOR source.

Level 0	Level 1	Level 2	Level 3
Source	Source (Operations)	SS1	SS1.1, SS1.2, SS1.3, SS1.4, SS1.5
		SS2	SS2.1, SS2.2, SS2.3, SS2.4, SS2.5
		SS3	SS3.1, SS3.2, SS3.3, SS3.4, SS3.5, SS3.6, SS3.7
	Source (Manage)	SES	SES1 through to SES10

Table 26: Axes structure for SCOR make.

Level 0	Level 1	Level 2	Level 3
Make	Make(Operations)	SM1	SM1.1, SM1.2, SM1.3, SM1.4, SM1.5, SM1.6, SM1.7
		SM2	SM2.1, SM2.2, SM2.3, SM2.4, SM2.5, SM2.6, SM2.7
		SM3	SM3.1, SM3.2, SM3.3, SM3.4, SM3.5, SM3.6, SM3.7, SM3.8
	Make (Manage)	SEM	SEM1 through to SEM9

Table 27: Axes structure for SCOR deliver.

Level 0	Level 1	Level 2	Level 3
Deliver	Deliver (Operations)	SD1	SD1.1, SD1.2, SD1.3, SD1.4, SD1.5, SD1.6, SD1.7, SD1.8, SD1.9, SD1.10, SD1.11, SD1.12, SD1.13, SD1.14, SD1.15
		SD2	SD2.1, SD2.2, SD2.3, SD2.4, SD2.5, SD2.6, SD2.7, SD2.8, SD2.9, SD2.10, SD2.11, SD2.12, SD2.13, SD2.14, SD2.15
		SD3	SD3.1, SD3.2, SD3.3, SD3.4, SD3.5, SD3.6, SD3.7, SD3.8, SD3.9, SD3.10, SD3.11, SD3.12, SD3.13, SD3.14, SD3.15
		SD4	SD4.1, SD4.2, SD4.3, SD4.4, SD4.5, SD4.6, SD4.7
	Deliver (Manage)	SEP	SED1 through to SED9

Table 28: Axes structure for SCOR return.

Level 0	Level 1	Level 2	Level 3
Return	Return (Operations)	SR1	SR1.1, SR1.2, SR1.3, SR1.4, SR1.5
		DR1	DR1.1, DR1.2, DR1.3, DR1.4
		SR2	SR2.1, SR2.2, SR2.3, SR2.4, SR2.5
		DR2	DR2.1, DR2.2, DR2.3, DR2.4
		SR3	SR3.1, SR3.2, SR3.3, SR3.4, SR3.5
		DR3	DR3.1, DR3.2, DR3.3, DR3.4
	Return (Manage)	SEP	SER1 through to SER9

Each level 3 process in the structure is connected to one or more other process by a link. These links are either inter- process or intra process, for instance, process SS1.1 links, by inter-process, to SS1.2. It also links, by intra-process, to SP2.1, SEP10, SS3.4, SES10, SM1.1, SM2.1, SM3.2, SD1.2, SD1.8, and SD4.2. Process SS1.1 also has a link to a process outside the SCOR^{TM} model, hence a category ‘Non Graph Out’ (NGO) has been created to capture links to processes outside the SCOR^{TM} model. The process SS1.1 therefore has an out degree k of 12, the sum of all the links.

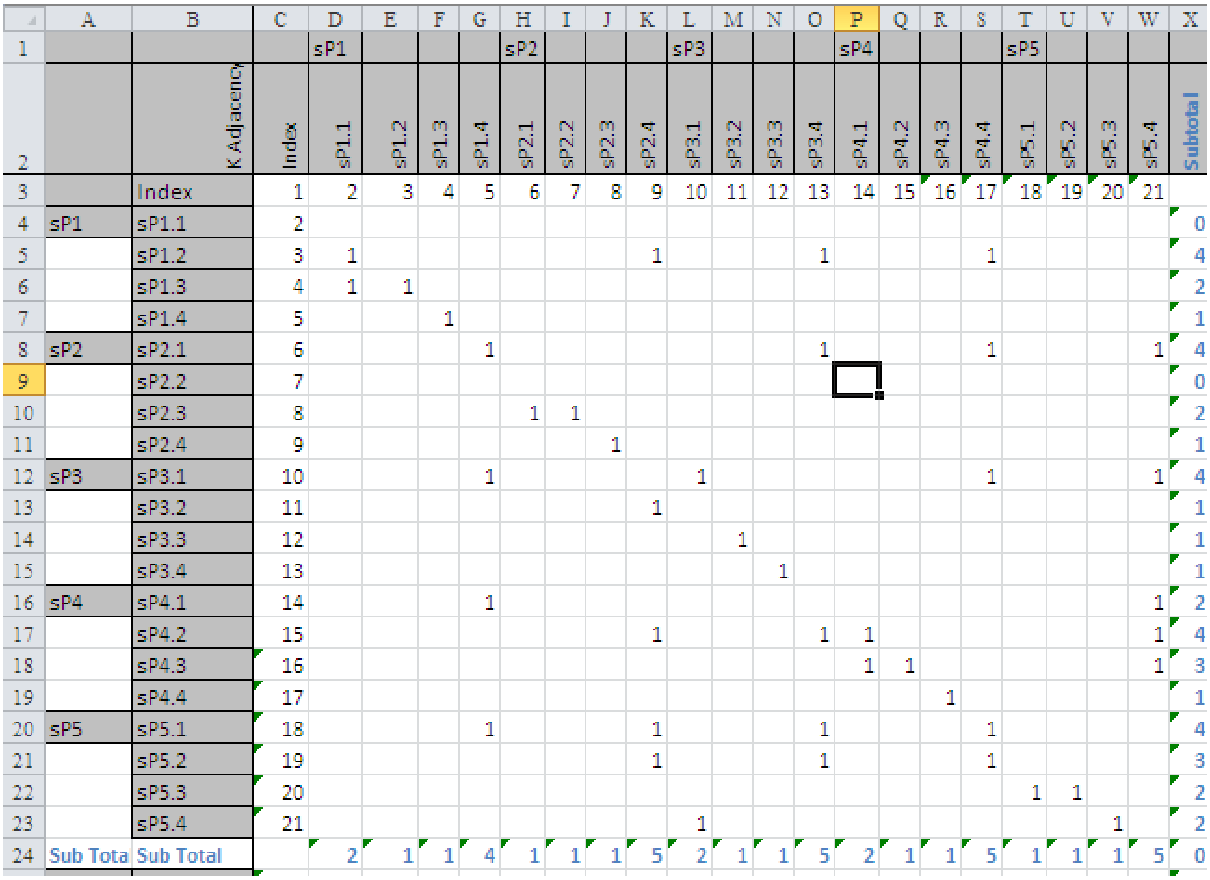
In chapter two the difference between bi-directional and uni-directional links was explained. The SCOR^{TM} model consists solely of uni-directional links; as a consequence the adjacency matrix is asymmetrical about the diagonal. The same section also explained the concept of circularity – where a link loops back on itself to the originating process. In the SCOR^{TM} model such loops are rare, but they do exist. In such cases these can be seen as values in the adjacency matrix diagonal; for instance, process SD1.14 has a link that loops back to the process.

As a consequence of the structure – the differing number of activities in the processes and the differing numbers of connectors, represented as nodes and links - the number of variables in each SCOR element differs. Figure 29 sets out the number of variables (in brackets) at each level.

Table 29: Aggregation structure in the SCOR Model.

	Plan	Source	Make	Deliver	Return
Level Three	SP1.1 to SEP10 (30)	SS1.1 to SES10 (27)	SM1.1 to SEM9 (31)	SD1.1 to SED9 (61)	SR1.1 to SER9 (36)
Level Two	SP1 to SEP10 (15)	SS1 to SES10 (13)	SM1 to SEM9 (12)	SD1 to SED9 (13)	SR1 to SER9 (15)
Level One	SP (Ops) and SEP (Manage) (2)	SS (Ops) and SES (Manage) (2)	SM (Ops) and SEM (Manage) (2)	SD (Ops) and SED (Manage) (2)	SR (Ops) and SER (Manage) (2)
Level Zero	All planning (1)	All sourcing (1)	All make (1)	All deliver (1)	All return (1)

Microsoft excel, with the NodeXL add in has been used as an analysis tool. All values in the spreadsheet have a limit of five decimal places and all values in the tables are taken directly from the spreadsheets; consequently, adding values from the tables may show a small error from the summation value in the table due to the rounding process of the spreadsheet. A sample of the spread sheet is shown in Figure 29 (Note the example of the circular reference in cell L12):



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1				sP1				sP2				sP3				sP4				sP5				
2		K Adjacency	Index	sP1.1	sP1.2	sP1.3	sP1.4	sP2.1	sP2.2	sP2.3	sP2.4	sP3.1	sP3.2	sP3.3	sP3.4	sP4.1	sP4.2	sP4.3	sP4.4	sP5.1	sP5.2	sP5.3	sP5.4	Subtotal
3		Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
4	sP1	sP1.1	2																					0
5		sP1.2	3	1							1				1				1					4
6		sP1.3	4	1	1																			2
7		sP1.4	5			1																		1
8	sP2	sP2.1	6				1								1				1				1	4
9		sP2.2	7																					0
10		sP2.3	8					1	1															2
11		sP2.4	9							1														1
12	sP3	sP3.1	10				1					1							1				1	4
13		sP3.2	11								1													1
14		sP3.3	12										1											1
15		sP3.4	13											1										1
16	sP4	sP4.1	14				1																1	2
17		sP4.2	15								1				1	1							1	4
18		sP4.3	16													1	1						1	3
19		sP4.4	17															1						1
20	sP5	sP5.1	18				1				1				1				1					4
21		sP5.2	19								1				1				1					3
22		sP5.3	20																	1	1			2
23		sP5.4	21									1										1		2
24	Sub Tota	Sub Total		2	1	1	4	1	1	1	5	2	1	1	5	2	1	1	5	1	1	1	5	0

Figure 29: Sample of the SCOR adjacency matrix in excel

The aim of this section was to set out the basic construction of the adjacency matrix. In addition to a basic analysis, the methodology calls for the analysis to be completed using two business scenarios; the prime and sub-contract scenarios. The next section sets out the construction of these two scenarios.

5.1.2 Construction of the Business Scenarios

In chapter four the methodology for analysis was explained. The section proposed a comparison of entropy – including the hierarchical construct necessary when aggregation is used - with the hidden information approach proposed in this research. In addition to the standard SCOR process model the methodology proposed field research to identify two further scenarios representing typical business paradigms. Sixteen structured interviews were completed using the questionnaire set out in appendix D; one for each of the two scenarios. The scenario for the creation of a subset scenario will be considered first; followed by the creation of the superset scenario.

5.1.2.1 Construction of the subset scenario

For simplicity, the structured questioning gave the interviewee the opportunity to select the components of the scenario at any level in the SCOR model. Thus, for instance, if the interviewee considered a scenario to include all the ‘plan’ section, they could select this section at level zero. Alternatively, if the interviewee considered a scenario to need a more detailed definition, this could be completed by selecting the appropriate processes at the appropriate level. Where an interviewee selects a section or sub-section, all processes within that section are assumed to have been selected. The outcome of this selection process identifies the sections and processes thought, by the interviewees, to be representative of the scenario. Once identified, these processes will be defined as representative of this scenario..

Sixteen responses were completed in the allotted timescale; of the sixteen, eight chose a scenario reflecting this sub-contractor scenario. For this scenario the results are set out in Table 30.

Table 30: Sub-contractor scenario process decisions.

Respondent	All Planning Processes	All Source Processes	All Make Processes	All Deliver Processes	All Return Processes
a	Yes	Yes	Yes	Yes	Yes
b	No	No	Yes	Yes	Yes
c	No	Yes	Yes	Yes	Yes
d	No	Yes	Yes	Yes	Yes
e	No	Yes	Yes	Yes	No
f	No	No	Yes	Yes	Yes
g	No	No	Yes	Yes	Yes
h	No	No	Yes	Yes	Yes
Yes Count	1	4	8	8	7
No Count	7	4	0	0	1

For this scenario the interviewees selected scenarios at three levels. The number of interviewees selecting each level is shown in Table 31:

Table 31: The number of interviewees selecting scenarios for each level in the model.

	Level Zero	Level One	Level Two	Level Three
Plan		1	0	7
Source		4	1	3
Make		8		
Deliver		8		
Return		7	1*	

*Note. One respondent selected level two responses but exited the structured questions before completion of the complete question set at level two. Never the less, the assumption herein is that the level two processes were, in the view of that respondent, at the necessary and sufficient level of granularity.

The methodology stated Fleiss kappa (F_k) to be the determinant of the reliability of agreement among the contributors:

$$F_k = \frac{\bar{P} - \bar{\bar{P}}_e}{1 - \bar{P}_e} = \frac{0.786 - 0.58}{1 - 0.58}$$

Equation 56

$$= 0.490$$

where

$$\bar{P} = \tfrac{1}{N} \textstyle\sum_1^N P_i = 0.786$$

Equation 57

and

$$\bar{P}_e = \sum_1^k p_j^2 = 0.58$$

Equation 58

Fleiss kappa, with a value of 0.490 suggests only a moderate reliability of an agreement that all plan, source, make, deliver and return processes would be part of the sub-contractor scenario. Further analysis of Table 30 shows there to be a number of consistencies and inconsistencies. Firstly there is consistency in the view that the planning processes would not be included in the sub-contractor scenario. Secondly, there is consistency in the view that the make, deliver and return processes would be included in this scenario. There is inconsistency in the view that the source process would be included in the scenario. Thus, from the field research results and for the sub-contractor scenario all the plan and source processes will be excluded from the analysis; all the make, deliver and return processes will be included in the analysis. On this basis, Table 32 to Table 36 set out the processes constituting the set of processes forming the sub-contractor scenario.

Table 32: Axes structure for SCOR plan.

Level 0	Level 1	Level 2	Level 3 Processes	Level 3 status for sub-contractor scenario
Plan	Plan (Operations)	SP1	SP1.1, SP1.2, SP1.3, SP1.4	All processes excluded
		SP2	SP2.1, SP2.2, SP2.3, SP2.4	All processes excluded
		SP3	SP3.1, SP3.2, SP3.3, SP3.4	All processes excluded
		SP4	SP4.1, SP4.2, SP4.3, SP4.4	All processes excluded
		SP5	SP5.1, SP5.2, SP5.3, SP5.4	All processes excluded
	Plan (Manage)	SEP	SEP1 through to SEP10	All processes excluded

Table 33: Axes structure for SCOR source.

Level 0	Level 1	Level 2	Level 3 Included	Level 3 status for sub-contractor scenario
Source	Source (Operations)	SS1	SS1.1, SS1.2, SS1.3, SS1.4, SS1.5	All processes excluded
		SS2	SS2.1, SS2.2, SS2.3, SS2.4, SS2.5	All processes excluded
		SS3	SS3.1, SS3.2, SS3.3, SS3.4, SS3.5, SS3.6, SS3.7	All processes excluded
	Source (Manage)	SES	SES1 through to SES10	All processes excluded

Table 34: Axes structure for SCOR make.

Level 0	Level 1	Level 2	Level 3 Included	Level 3 status for sub-contractor scenario
Make	Make(Operations)	SM1	SM1.1, SM1.2, SM1.3, SM1.4, SM1.5, SM1.6, SM1.7	All processes included
		SM2	SM2.1, SM2.2, SM2.3, SM2.4, SM2.5, SM2.6, SM2.7	All processes included
		SM3	SM3.1, SM3.2, SM3.3, SM3.4, SM3.5, SM3.6, SM3.7, SM3.8	All processes included
	Make (Manage)	SEM	SEM1 through to SEM9	All processes included

Table 35: Axes structure for SCOR deliver.

Level 0	Level 1	Level 2	Level 3 Included	Level 3 status for sub-contractor scenario
Deliver	Deliver (Operations)	SD1	SD1.1, SD1.2, SD1.3, SD1.4, SD1.5, SD1.6, SD1.7, SD1.8, SD1.9, SD1.10, SD1.11, SD1.12, SD1.13, SD1.14, SD1.15	All processes included
		SD2	SD2.1, SD2.2, SD2.3, SD2.4, SD2.5, SD2.6, SD2.7, SD2.8, SD2.9, SD2.10, SD2.11, SD2.12, SD2.13, SD2.14, SD2.15	All processes included
		SD3	SD3.1, SD3.2, SD3.3, SD3.4, SD3.5, SD3.6, SD3.7, SD3.8, SD3.9, SD3.10, SD3.11, SD3.12, SD3.13, SD3.14, SD3.15	All processes included
		SD4	SD4.1, SD4.2, SD4.3, SD4.4, SD4.5, SD4.6, SD4.7	All processes included
	Deliver (Manage)	SEP	SED1 through to SED9	All processes included

Table 36: Axes structure for SCOR return.

Level 0	Level 1	Level 2	Level 3 Included	Level 3 status for sub-contractor scenario
Return	Return (Operations)	SR1	SR1.1, SR1.2, SR1.3, SR1.4, SR1.5	All processes included
		DR1	DR1.1, DR1.2, DR1.3, DR1.4	All processes included
		SR2	SR2.1, SR2.2, SR2.3, SR2.4, SR2.5	All processes included
		DR2	DR2.1, DR2.2, DR2.3, DR2.4	All processes included
		SR3	SR3.1, SR3.2, SR3.3, SR3.4, SR3.5	All processes included
		DR3	DR3.1, DR3.2, DR3.3, DR3.4	All processes included
	Return (Manage)	SEP	SER1 through to SER8	All processes included

Given the above, in this subset scenario, the process model is representative of a business acting as a subcontractor in a business transaction; that is, where planning and sourcing are completed by the prime contractor, with the business model for the subcontractor being limited to make, deliver and return from the SCOR process model. This subcontractor position is, obviously, just one side of the business transaction. The opposite side of the same transaction would be representative of the buying organisation, the prime contractor. The next section will, if possible, determine the process model required for an organisation to act in such a role.

5.1.2.2 Construction of the superset scenario

For the construction of this scenario, and again for simplicity, the approach gave the interviewee the opportunity to select the sections of the scenario at any level in the SCOR model; thus again, for instance, if the interviewee considered a scenario to include all the ‘plan’ section, they could select this section at level zero. Alternatively, if the interviewee considered a scenario to need a more detailed definition, this could be completed by selecting the appropriate processes at the appropriate level. The key difference between the construction of this scenario and that of the sub-contractor scenario above is that, in this case, the interviewee is asked to identify only the processes that are additional to the SCOR model, as would be necessary when operating in the prime contracting buyer role in the transaction; i.e. the assumption is that the full SCOR model is required for the basic operation of the business and that elements of the model need to be duplicated for the prime contract role. For instance, the prime contractor may plan and for a major contract and again for the components of the contract for which it is directly responsible - at a lower level in the prime contract work breakdown structure. An example of this type of business structure can be seen in the automotive and pharmaceutical industries. In the case of an automotive manufacturer, if the engine is a bought out component, the output of one planning activity is demand onto the engine supplier. In the scenario where the engine is built by the same organisation with a make to stock buffer, planning is done once for the vehicle plan and once again for the engine plan; thus, in this scenario, the planning capability is executed twice. In the pharmaceutical industry planning and sourcing are completed initially by the business operating unit (usually a geographically based unit), with production being planned and sourced (usually to a contract manufacturing organisation) a second time by the production operating unit.

Sixteen responses were completed in the allotted timescale; of the sixteen, eight chose a scenario reflecting this prime contractor scenario. For this scenario the results are set out in Table 37.

Table 37: Sub-contractor scenario process decisions.

Respondent	All Planning Processes	All Source Processes	All Make Processes	All Deliver Processes	All Return Processes
i	Yes	Yes	Yes	Yes	No
j	Yes	Yes	No	No	No
k	Yes	Yes	No	No	No
l	Yes	Yes	Yes	Yes	Yes
m	Yes	Yes	No	No	No
n	Yes	Yes	No	No	Yes
o	Yes	Yes	Yes	Yes	Yes
p	Yes	Yes	No	No	No
Yes Count	8	8	3	3	3
No Count	0	0	5	5	5

In this scenario the interviewees selected scenarios at the levels shown in Table 38:

Table 38: The number of interviewees selecting scenarios for each level in the model.

	Level Zero	Level One	Level Two	Level Three
Plan		8		
Source		8		
Make		3		5
Deliver		3	2	3
Return		3	3	2

Where an interviewee selects a section, all processes within that section are assumed to have been selected. The outcome of this selection process identifies the sections and processes thought, by the interviewees, to be representative of the scenario. Once identified, these processes will be defined as representative of this scenario.

The methodology stated Fleiss kappa (F_k) to be the determinant of the reliability of agreement among the contributors:

$$F_k = \frac{\bar{P} - \bar{\bar{P}}_e}{1 - \bar{P}_e} = \frac{0.679 - 0.531}{1 - 0.531}$$

Equation 59

$$= 0.314$$

where

$$\bar{P} = \tfrac{1}{N}\textstyle\sum_1^N P_i = 0.679$$

Equation 60

and

$$\bar{P}_e = \sum_1^k p_j^2 = 0.531$$

Equation 61

Fleiss kappa, with a value of 0.314 suggests only a moderate reliability of an agreement that all plan, source, make, deliver and return processes would be part of the prime-contractor scenario. Analysing Table 37 in more detail – as was done in the sub-contractor scenario - we can see there are a number of consistencies and inconsistencies. Firstly there is consistency in the view that the planning and sourcing processes would, for part of the additional process set, be required in the prime-contractor scenario. There is inconsistency in the view that the make, deliver and return process would be included in the scenario as additional processes. Thus, from the field research results and for the prime-contractor scenario, all the plan and source processes – where there is a high reliability of agreement - will be included as additional processes in the set of processes representing the prime-contractor scenario. Make, deliver and return processes will be excluded from the set of additional processes in this case.

On this basis, Table 39 to Table 43 set out the processes constituting the set of processes forming the prime contractor scenario.

Table 39: Axes structure for SCOR plan.

Level 0	Level 1	Level 2	Level 3 Included	Level 3 status for prime-contractor scenario
Plan	Plan (Operations)	SP1	SP1.1, SP1.2, SP1.3, SP1.4	All processes are included
		SP2	SP2.1, SP2.2, SP2.3, SP2.4	All processes are included
		SP3	SP3.1, SP3.2, SP3.3, SP3.4	All processes are included
		SP4	SP4.1, SP4.2, SP4.3, SP4.4	All processes are included
		SP5	SP5.1, SP5.2, SP5.3, SP5.4	All processes are included
	Plan (Manage)	SEP	SEP1 through to SEP10	All processes are included

Table 40: Axes structure for SCOR source.

Level 0	Level 1	Level 2	Level 3 Included	Level 3 status for prime-contractor scenario
Source	Source (Operations)	SS1	SS1.1, SS1.2, SS1.3, SS1.4, SS1.5	All processes are included
		SS2	SS2.1, SS2.2, SS2.3, SS2.4, SS2.5	All processes are included
		SS3	SS3.1, SS3.2, SS3.3, SS3.4, SS3.5, SS3.6, SS3.7	All processes are included
	Source (Manage)	SES	SES1 through to SES10	All processes are included

The processes set out in Table 37 and Table 40 will be included twice to represent the duality of the plan and source process set.

Table 41: Axes structure for SCOR make.

Level 0	Level 1	Level 2	Level 3 Included	Level 3 status for prime-contractor scenario
Make	Make (Operations)	SM1	SM1.1, SM1.2, SM1.3, SM1.4, SM1.5, SM1.6, SM1.7	All processes are excluded from the additional set.
		SM2	SM2.1, SM2.2, SM2.3, SM2.4, SM2.5, SM2.6, SM2.7	All processes are excluded from the additional set.
		SM3	SM3.1, SM3.2, SM3.3, SM3.4, SM3.5, SM3.6, SM3.7, SM3.8	All processes are excluded from the additional set.
	Make (Manage)	SEM	SEM1 through to SEM9	All processes are excluded from the additional set.

Table 42: Axes structure for SCOR deliver.

Level 0	Level 1	Level 2	Level 3 Included	Level 3 status for prime-contractor scenario
Deliver	Deliver (Operations)	SD1	SD1.1, SD1.2, SD1.3, SD1.4, SD1.5, SD1.6, SD1.7, SD1.8, SD1.9, SD1.10, SD1.11, SD1.12, SD1.13, SD1.14, SD1.15	All processes are excluded from the additional set.
		SD2	SD2.1, SD2.2, SD2.3, SD2.4, SD2.5, SD2.6, SD2.7, SD2.8, SD2.9, SD2.10, SD2.11, SD2.12, SD2.13, SD2.14, SD2.15	All processes are excluded from the additional set.
		SD3	SD3.1, SD3.2, SD3.3, SD3.4, SD3.5, SD3.6, SD3.7, SD3.8, SD3.9, SD3.10, SD3.11, SD3.12, SD3.13, SD3.14, SD3.15	All processes are excluded from the additional set.
		SD4	SD4.1, SD4.2, SD4.3, SD4.4, SD4.5, SD4.6, SD4.7	All processes are excluded from the additional set.
	Deliver (Manage)	SEP	SED1 through to SED9	All processes are excluded from the additional set.

Table 43: Axes structure for SCOR return.

Level 0	Level 1	Level 2	Level 3 Included	Level 3 Excluded
Return	Return (Operations)	SR1	SR1.1, SR1.2, SR1.3, SR1.4, SR1.5	All processes are excluded from the additional set.
		DR1	DR1.1, DR1.2, DR1.3, DR1.4	All processes are excluded from the additional set.
		SR2	SR2.1, SR2.2, SR2.3, SR2.4, SR2.5	All processes are excluded from the additional set.
		DR2	DR2.1, DR2.2, DR2.3, DR2.4	All processes are excluded from the additional set.
		SR3	SR3.1, SR3.2, SR3.3, SR3.4, SR3.5	All processes are excluded from the additional set.
		DR3	DR3.1, DR3.2, DR3.3, DR3.4	All processes are excluded from the additional set.
	Return (Manage)	SEP	SER1 through to SER8	All processes are excluded from the additional set.

In this superset scenario the business model is representative of a business acting as a prime contractor in a business transaction, where planning and sourcing are completed by the buying organisation in the role of prime contractor; that is, they are planning and sourcing at a ‘project’ or ‘programme’ level with their own internal organisation, and the sub-contractor organisations carrying out work in response to the plan. With the two scenarios – prime and subcontractors – this research covers a basic construct of both the buyer or prime contractor, and seller or subcontractor sides of a capability asymmetric transaction.

This section has explained the construction of the basic adjacency matrix and the content of the subset and superset scenarios to be applied to the adjacency matrix. The results of the analysis will cover the horizontal/macro analysis and the vertical/micro analysis for the basic, subset and superset scenarios.

For each level of analysis the research will compare the results from the entropy method with those of the hidden information model proposed herein; each scenario will be analysed separately. Prior to commencing an analysis, and to maintain construct validity, an explanation of the construction of the analysis worksheets is provided; consequently, as a start to the basic analysis this construction is explained.

5.2 Analysis of the basic scenarios

5.2.1 Analysis of the Basic Construct.

The outcome of the construction is a 186 by 186 adjacency matrix (Figure 29 shows the first 24 by 24 entries). Initially the results should validate the construction of the adjacency matrix, from the supply chain operating reference model, aligns with the characteristics and measures of network analysis.

Excluding links to external processes (Non Graph Out links, [NGO]), the process degree distribution for K_{out} , for the complete adjacency matrix, ranges between a minimum of zero and a maximum of twenty six; this being process sM3.4 (Produce and test, in the Make subsection). The mean number of links is 6.18919 recurring, and the median is 5. The distribution of process K_{out} degrees for the complete matrix is shown in Figure 30.

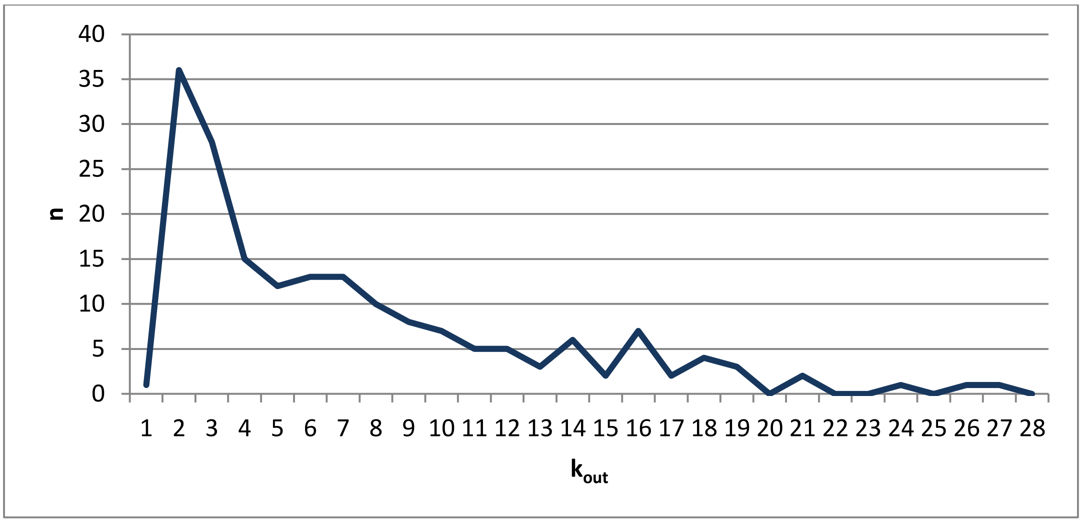


Figure 30: Graph of Kout degree frequencies

By comparison, K_{in} for the complete adjacency matrix ranges from a minimum of zero to a maximum of twenty three; this being process sP5.1 (Assess and aggregate returns requirements, within the Plan subsection). The mean number of links is 6.38919 and the median value is 4. The frequencies of process K_{in} degree values are shown in Figure 31.

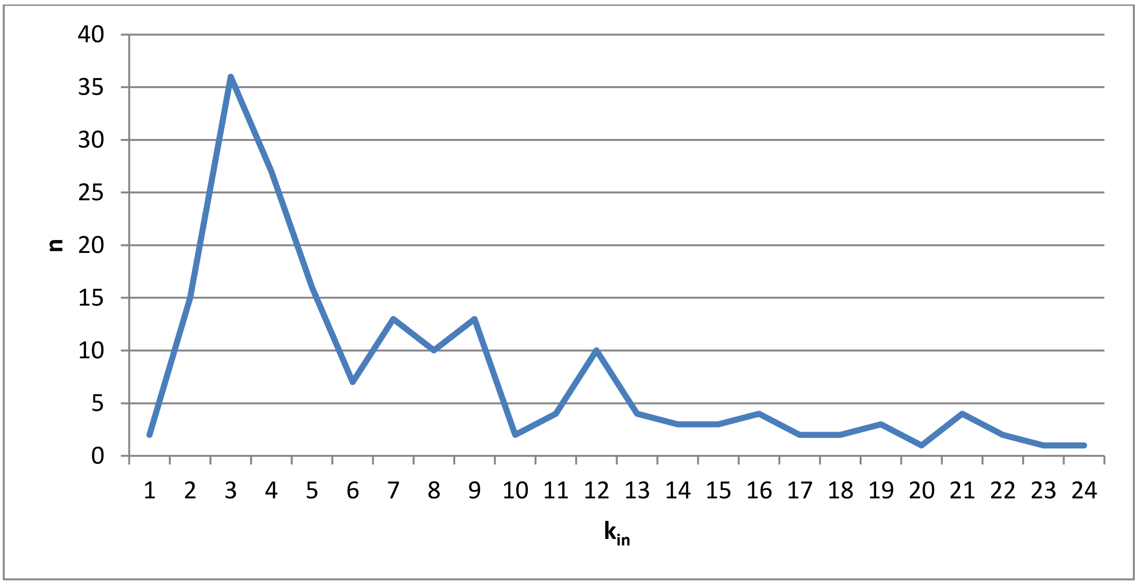


Figure 31: Graph of Kin degree frequencies

The K_{out} minimum, maximum, mean and median, for the complete matrix can be broken down into the plan, source, make, deliver and return process subsections, represented at level three in the SCOR model. The breakdown, into the level one values, is shown in Table 44.

Table 44: Table of values for the level three processes.

	Plan	Source	Make	Deliver	Return
Max	20	20	26	17	18
Min	1	1	1	0	1
Mean	5.06667	8.88889	7.29032	4.16393	7.58333
Median	2.5	9	5	2	7

This type of degree distribution is classic in network analysis; for comparison, Newman (2010) offers degree distribution values for the world wide web; reproduced in Figure 32:

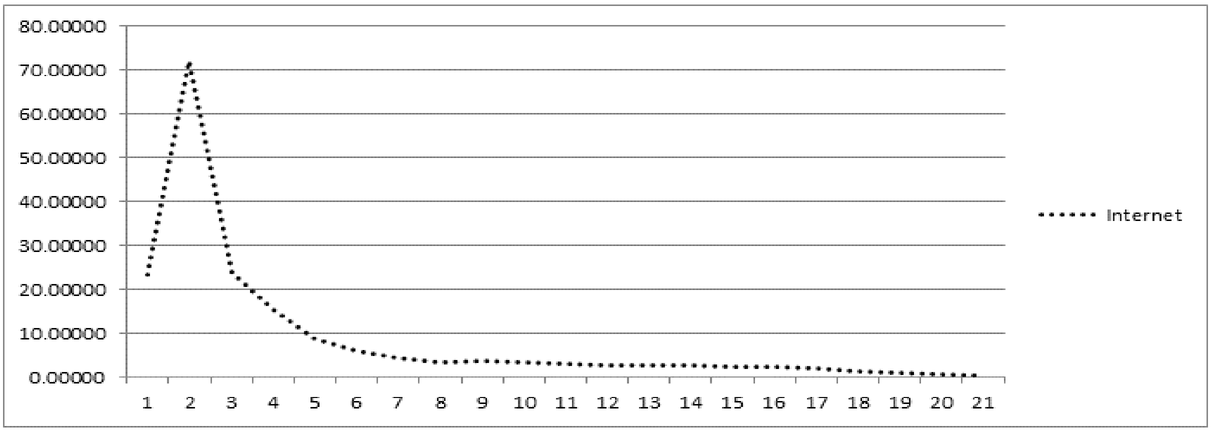
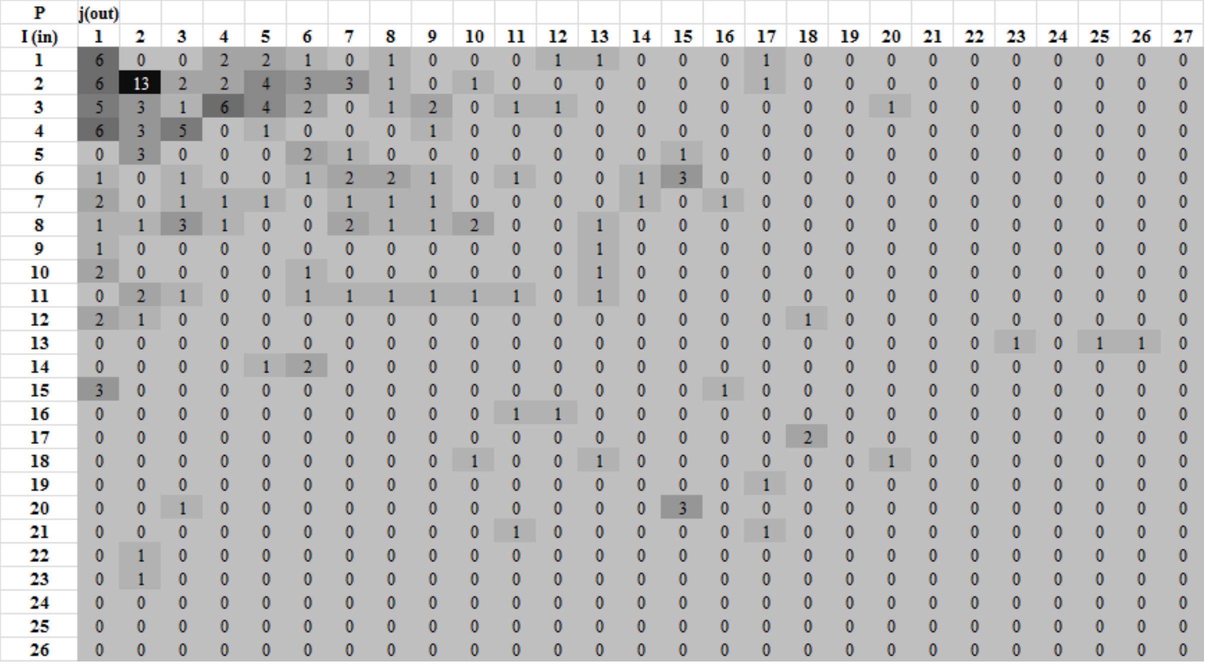


Figure 32: In degree distribution from the world wide web (adapted from Newman (2010)).

Assuming an alternative hypothesis to be there is no statistically significant correlation between the in degree of the internet and that of the adjacency matrix derived here, and the null hypothesis to be that a statistically significant correlation exists. Pearson product moment correlation co-efficient $r_{obt} = 0.41$ compares to a two tailed test value for alpha = 0.5 $r_{cv} = \pm 0.4438$. Thus $r(18) = +0.41, p > .05\ (two - tailed)$.

This shows the classic decay structure, prevalent in networks, applies to the adjacency matrix created from the supply chain operations reference model. Newman (2010) further points out the lack of research analysis of joint in degree out degree correlation and an understanding of the correspondence in the decay function for each degree; that is, where a specific value for k_{in} corresponds to a specific value for k_{out} Figure 31 shows the joint degree correlation for the this adjacency matrix.



P	j(out)																										
I (in)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**
1	6	0	0	2	2	1	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
2	6	13	2	2	4	3	3	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
3	5	3	1	6	4	2	0	1	2	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
4	6	3	5	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	3	0	0	0	2	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
6	1	0	1	0	0	1	2	2	1	0	1	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0
7	2	0	1	1	1	0	1	1	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
8	1	1	3	1	0	0	2	1	1	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	2	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	2	1	0	0	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0
14	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
20	0	0	1	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
22	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 33: Joint in/out degree for adjacency matrix

The table serves simply to demonstrates some correlation between in degree and out degree values of the adjacency matrix; that is, there is some correspondence in the decay of the distribution for the in and out degree values.

Another way to look at this distribution is by considering power law distribution, that is:

$$\log_2 p_k = \propto \log_2 k + c \qquad \text{Equation 62}$$

where $\propto$ and c are constants. This type of distribution is common in networks and is characterised as an approximation because the power law is non-monotonic for small k values and omits k values of zero. For this specific model, all k values are relatively small; never the less, the network does approximate to a power law distribution and demonstrates non-monotonicity for small k.

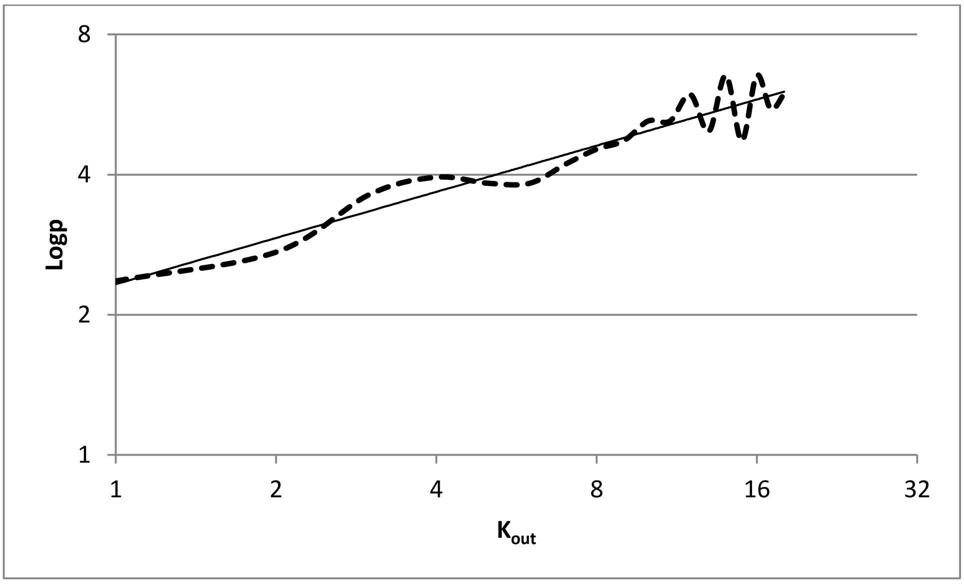


Figure 34: Matrix power law distribution

A third validating characteristic is clustering. The basics of clustering as a characteristics that gives rise to communities or graph components, the nodes of which may, or may not, be reachable from other nodes in the graph were explained in chapter two. The chapter also explained how, for directed graphs, the idea is more complex; breaking the concept down into three different descriptions. Firstly strongly connected components, used to describe a situation where there is a path AB and a corresponding path BA; thus, acyclic directed components cannot be described as strongly connected. Secondly, an out-component is a directed network the nodes of which are reachable from the origin (i.e. you can get out from the origin to the component, but you cannot get from the component to the origin). Finally, an in component graph is a directed network the origin of which is reachable from the nodes (i.e. you can get from the graph component to the origin, but not from the origin to the component). NodeXL provides measures of global clustering co-efficients which, for the adjacency matrix under analysis here, gives an average clustering coefficient value of 0.153, and a median value of 0.148, demonstrating a generally loosely coupled structure. Note: Loosely coupled structures suggest a low degree of connectivity with a significant number of structural holes in the matrix; which implies a low degree of connectivity in the network. Figure 35 to Figure 42, show a pictorial structure of the SCOR model. Figure 35 shows the model as a helical structure and Figure 36 as a first angle projection. Figure 37 to Figure 42 are the same first angle projection, with the links for each of the respective plan, source, make, deliver, return and NGO subsets highlighted.

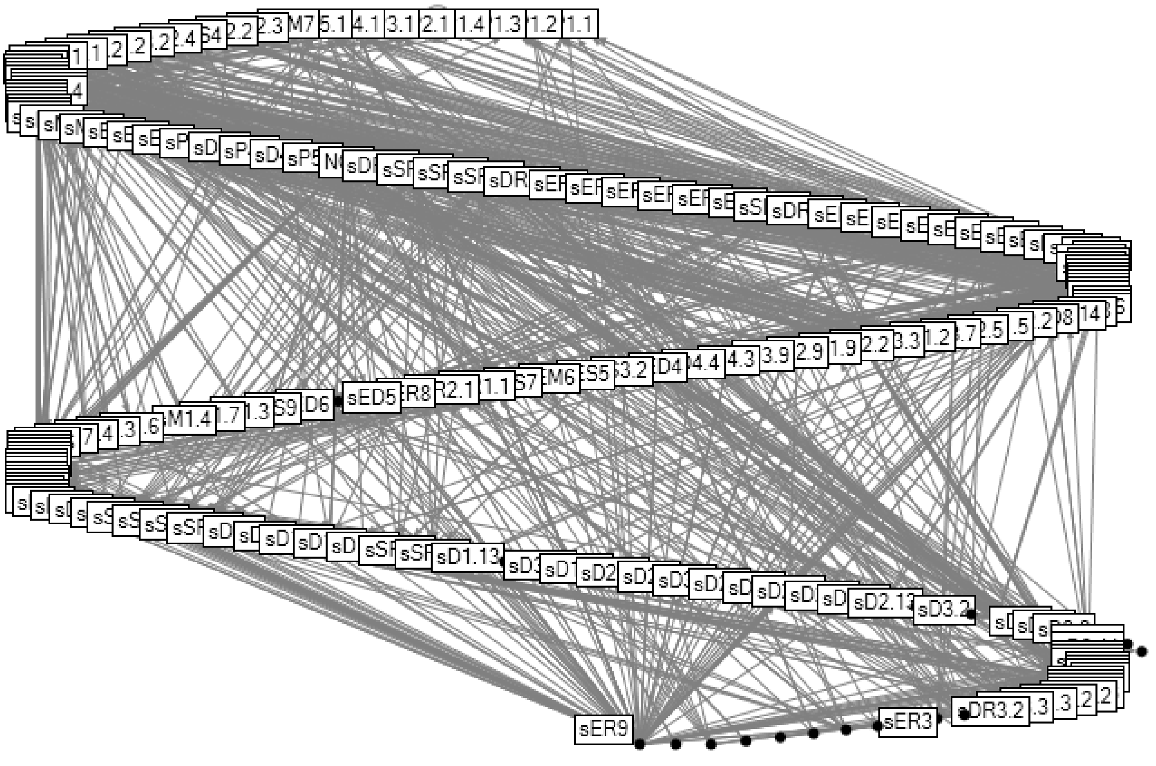


Figure 35: Visual map of the 186 nodes and 1261 links in the adjacency matrix.

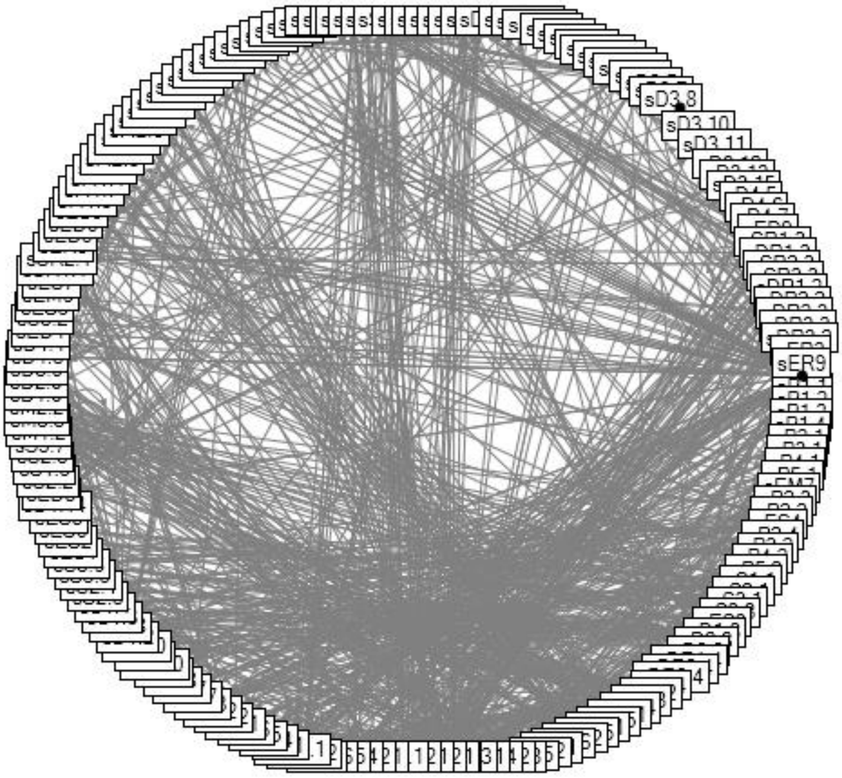


Figure 36: Circular visualisation (shows a view orthogonal to Figure 35)

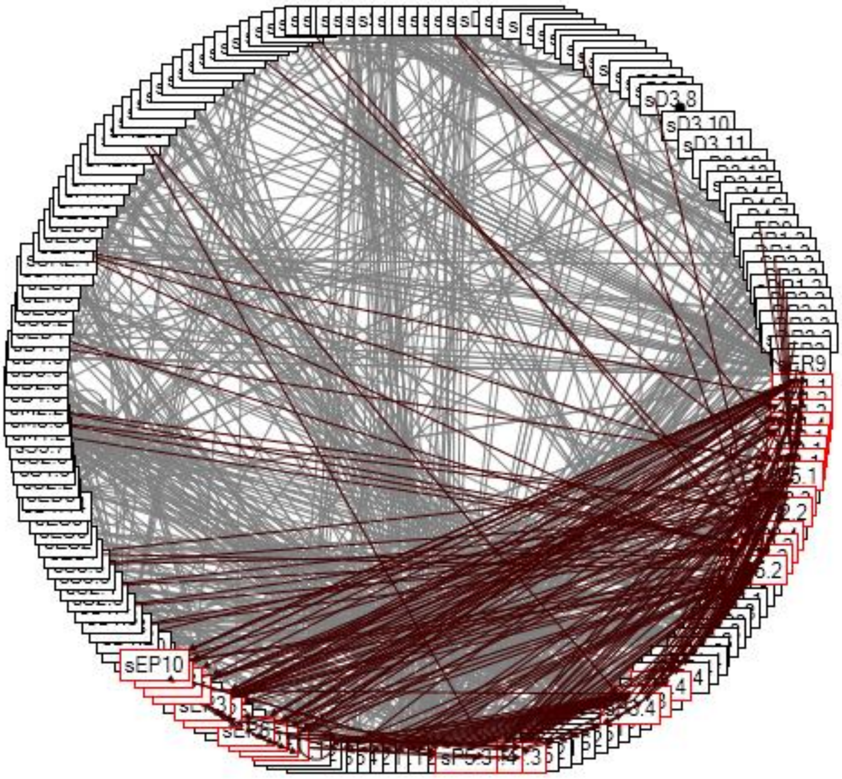


Figure 37: Plan links

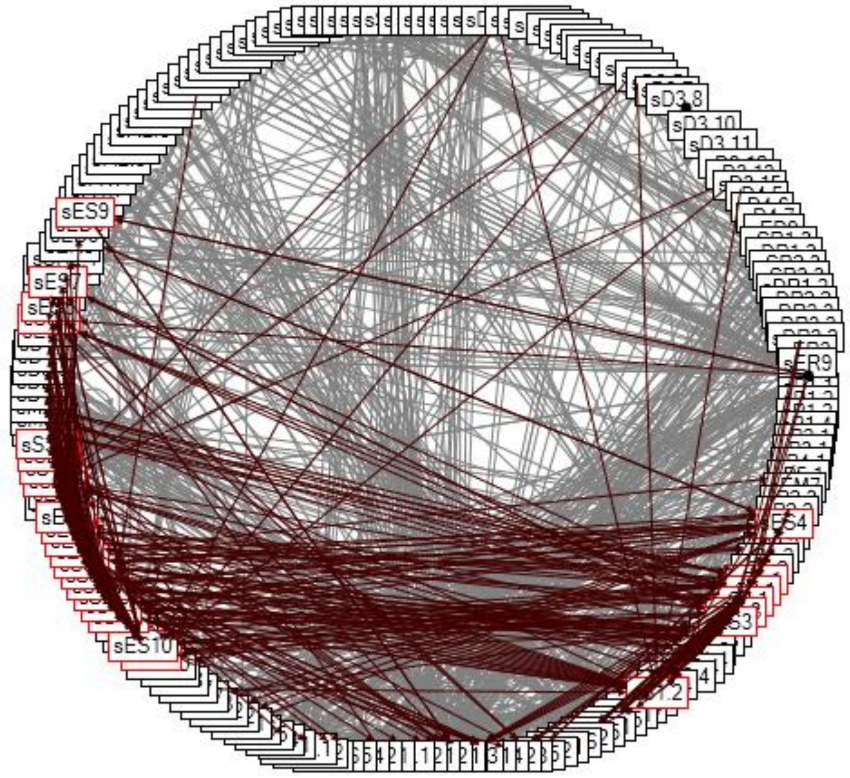


Figure 38: Source links

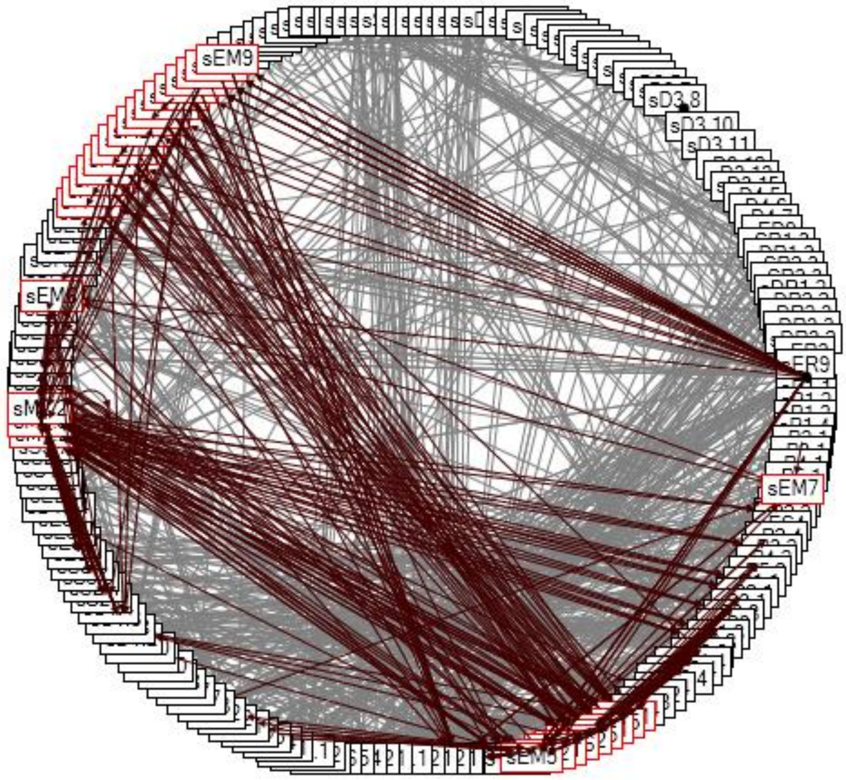


Figure 39: Make links

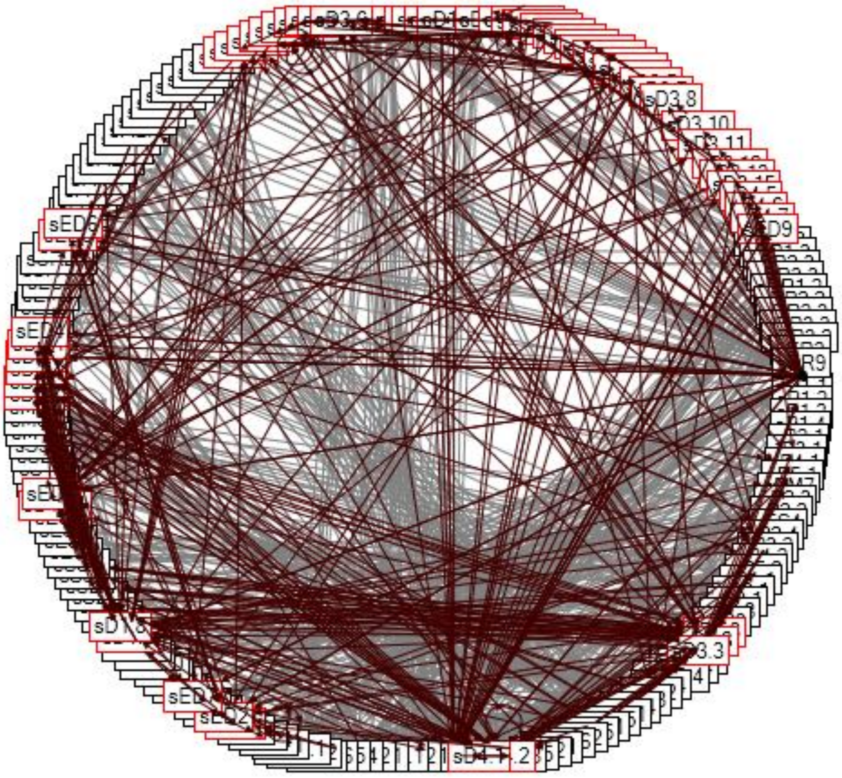


Figure 40: Deliver links

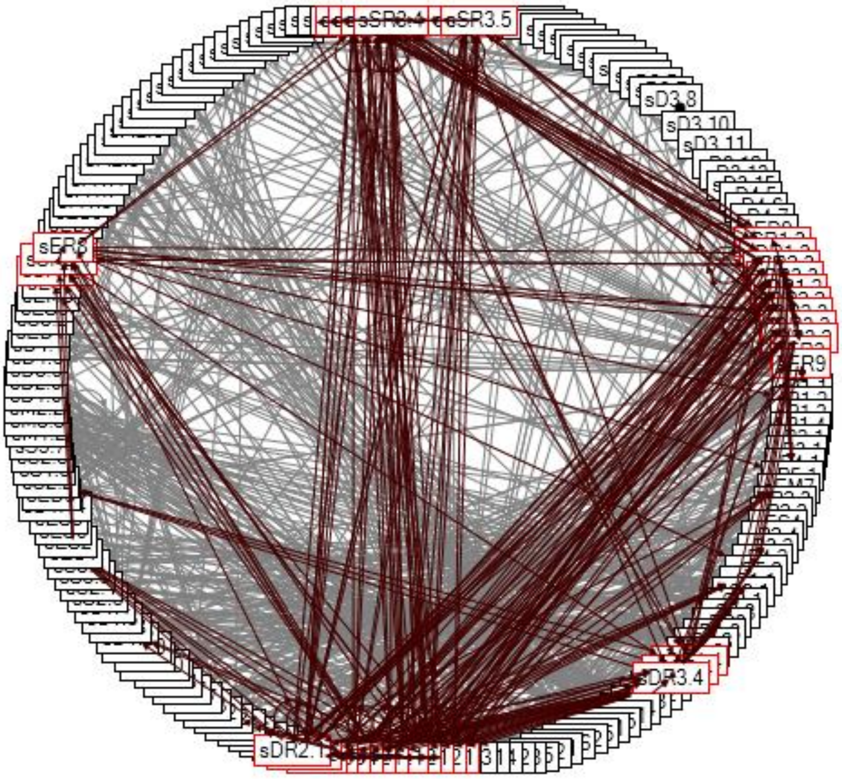


Figure 41: Return links

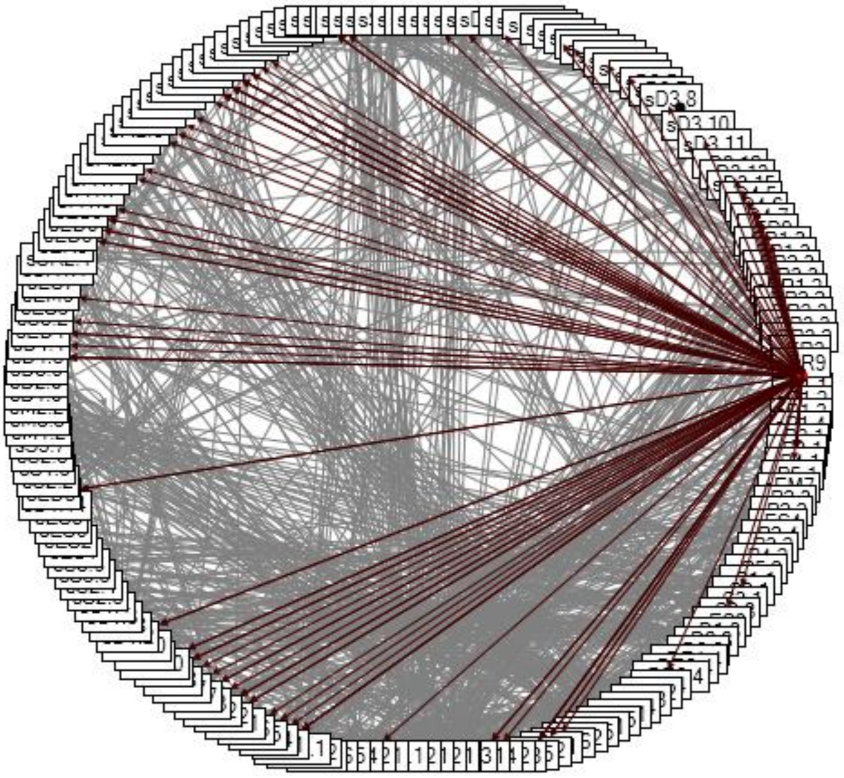


Figure 42: NGI links

The above provides a basic overview of the construction of the complete adjacency matrix for the supply chain operating reference model, the generally skewed distribution of $Kout$ (and $K\ in$) degree distribution, the maximum and minimum values, a comparison of out and in degree values, a breakdown of the out degree values by level one process and a pictorial view of the complete and segmented structures. However, what is of interest in this research is the horizontal/macro level and vertical/micro level comparison of entropy and hidden information, and how that informs insight into understanding the state of the business process model. The next section analyses these particular issues.

5.2.2 Macro Level Analysis for the Basic Construct

The theoretical foundation of entropy and information theory, in chapter two, explained the concept of baseline equi-probability where a baseline of equi-probable variables is constructed as the origin from where grouping or consolidation of data into information occurs.

Before any analysis is undertaken a baseline of equi-probability should be formed. For the SCOR model, using K_{out} degree as the independent variable for the measure of complexity, the baseline equi-probability (where all K are equal) equals:

$$S = \sum_{1}^{1261} p_i \log_2 p_i \ = \ \log_2 1261 \ = 10.3 \qquad \text{Equation 63}$$

where

$$p_i \ = \frac{1}{\sum_1^{186} k_{out}} \qquad \text{Equation 64}$$

That is, there are a total of 1261 links in the complete SCOR model, represented in the complete adjacency matrix. The complete matrix demonstrates an overall entropy measure of 6.81355, that is.

$$S = \sum_{1}^{186} p_i \log_2 p_i \ = \ 6.81355 \qquad \text{Equation 65}$$

where, in this case.

$$p_i \ = \frac{k_{iout}}{1261} \qquad \text{Equation 66}$$

That is, k_{iout} equals the out degree for the $i\ th$ SCOR process.

As shown in Table 45, the entropy value decreases to 5.19945 at the first level of aggregation; 3.20922 at the second level, and 2.52821 at the third level. Being additive, the overall entropy value is the sum of the plan, source, make, deliver, return and NGO for any given row. Thus, the structure of the SCOR model; the simple fact that k has differing probabilities, removes $10.3 - 6.81355 = 3.48645$ units of entropy through the structure of the SCOR model. That is, the allocation of each of the 1261 to one of the 186 processes removes the potential for some information to be contained within the SCOR model as the structure becomes more certain.

Table 45: Entropy for each of the SCOR Analysis Levels.

	Overall Entropy	Plan	Source	Make	Deliver	Return	NGO
Level Three	6.81355	0.87701 (30)	1.29262 (27)	1.23904 (31)	1.54758 (61)	1.54064 (36)	0.31677 (1)
Level Two	5.19945	0.78897 (15)	0.97651 (13)	0.89922 (12)	1.00376 (13)	1.21433 (15)	0.31677 (1)
Level One	3.20922	0.28104 0.20194	0.40198 0.20670	0.40518 0.16347	0.43776 0.14647	0.42977 0.21828	0.31667
Level Zero	2.5281	0.36794	0.45554	0.44450	0.46563	0.47793	0.31667

Considering Table 45, three further points can be identified. Firstly analysis at level zero includes six variables, one for each SCOR section and the NGO element. To ascertain the state of the system by differentiating between the six elements, and using binary question, would take $\log_2 6 = 2.58496$ questions and not 2.5281 as predicted by calculating entropy using the Shannon approach above. Practically, in the context of asking questions, this makes little difference; however, if the value is considered as a distance, that is, the distance in binary questions from an understanding of the state of the system, then the difference in the values is more important. Secondly, the fact that the value of entropy is not equal to $\log_2 6 = 2.58496$, suggests entropy to be measuring value other than a simple decision logic. Thirdly, the difference between the two values becomes greater with the level of granularity. At level three the values are 6.81355 and 7.5391 respectively, a more significant difference. Before this difference is considered further, a look at the same data from a hidden information perspective is required.

The hidden information analysis for the adjacency matrix, using the same overall structure, gives the values in Table 46. Chapter two explained how the formulation for hidden information is not, like entropy, an additive function; hence the constituent elements (plan, source, make, deliver, return and NGO) do not sum to the overall hidden information value as is the case for entropy in Table 45. For hidden information the values in the plan, source, make, deliver, return and NGO columns equate to $\log_2 n$ where n equals the number of values for each of these sections of the SCOR model.

Table 46: Hidden Information for the SCOR Analysis Levels.

	Overall Hidden Information	Plan	Source	Make	Deliver	Return	NGO
Level Three	7.53916	4.90689	4.75488	4.95419	5.93073	5.16992	0
Level Two	6.10852	3.90689	3.70044	3.58496	3.70044	3.90689	0
Level One	3.45943	1	1	1	1	1	0
Level Zero	2.58496	0	0	0	0	0	0

For information and comparison, Table 47 shows a weighted sum approach to arrive at an additive set of data for the hidden information approach. In this case the overall hidden value for the row shown equates to:

$$= \sum_{1}^{m} \frac{\log_2 n_m}{t} \qquad \text{Equation 67}$$

Where n equals the number of values in m groups and t equals the total number of values. So, for instance, using the level three return processes as an example:

$$5.16992 \times \frac{n_{m=Returns} = 36}{t = 187} = 0.99528 \qquad \text{Equation 68}$$

Table 47: Comparative weighted average values for level three.

	Overall Hidden Information	Plan	Source	Make	Deliver	Return	NGO
Level Three	5.22493	0.78720	0.68653	0.82128	1.93463	0.99528	0

Chapter two explained the problems with the use of entropy as anything other than a numerical representation of ‘mixed up ness’ applied at specific levels of analysis. It is here, at this level of analysis, that this phenomena is demonstrable in the data. Entropy, for the complete adjacency matrix (Table 45) has a value of 6.81355. This value has no unit of measure; i.e. it does not have a Boltzmann’s constant value k as applied in the statistical thermodynamic entropy variant (please do not confuse this value k , which represents Boltzmann’s constant in the statistical thermodynamic version of entropy, with $k_{out,}$ which represents the out degree of the adjacency matrix). Neither does, as has been demonstrated, the value equate to binary questioning. Using the overall entropy for level three – from Table 45 – as an example, it can been seen below that the outcome does not equate to the total number of determinable variables in the matrix.

$$2^{6.81355} = 112.48197$$ Equation 69

which is less than the 187 known variables in the complete adjacency matrix. To be pragmatic, and in reality, binary questioning would be rounded up to seven questions, this would still only allow for 128 variables to be analysed; 59 less than are necessary for the SCOR adjacency matrix. The same critique can be applied to entropy values at all levels of aggregation, even if these values are rounded up: At level two, rounded up values would give a maximum number of determinable values of 64 when there are 69 in the matrix. Level one, has 11 values in the matrix where entropy allows for $2^{3.20922} = 9.24850$ determinable values (rounding up to 4 would allow for 16 determinable values). Finally level zero, 6 values in the matrix where entropy would allow for $2^{2.52810} = 5.76811$ determinable values (note again that practical rounding up of questions would allow for 8 determinable values). Thus for the SCOR model, entropy, determined at each level of aggregation, is unlikely to add information on the state of the system except for providing an arbitrary measure of mixed up ness and only an approximation of the number of binary questions required. Being additive, the values of the subsets (Plan, Source, Make, Deliver, Return and NGO) offer very little in terms of understanding of the state of the subsystem. To analyse this in more detail, the next section looks at entropy and hidden information at each level in the SCOR structure; that is, at the plan, source, make, deliver and return levels.

Table 48 shows entropy for each distinct section of the SCOR model with the number of variables for each subset shown in brackets. To be clear, Table 45 considered the adjacency matrix as a whole, whereas Table 48 considers entropy for each section of the matrix. The reason for the difference in values between Table 48 and Table 45, for each of the SCOR sections is that in Table 45:

$$S = \sum_{1}^{187} p_i \log_2 p_i \qquad \text{Equation 70}$$

where

$$p_i = \frac{k_{iout}}{1261}$$ Equation 71

and for Table 48:

$$S = \sum_{1}^{187} p_i \log_2 p_i$$ Equation 72

where

$$p_i = \frac{\sum_1^i k_{mout}}{\sum_1^i k_{sout}}$$ Equation 73

Where m is the group and s is the section.

Table 48: Macro entropy for each distinct SCOR Level.

	Complete Matrix Entropy	Plan	Source	Make	Deliver	Return	NGO
Level Three	6.81355	4.22330 (30)	4.39818 (27)	4.43323 (31)	5.37138 (61)	4.90868 (36)	0.31677 (1)
Level Two	5.19945	3.49288 (15)	2.73726 (13)	2.3716 (12)	2.67157 (13)	3.40145 (15)	0.31677 (1)
Level One	3.20922	0.95443	0.80461	0.69269	0.58876	0.78566	0.31677
Level Zero	2.52821	0	0	0	0	0	0

For comparison, this can be viewed against the values for each section using the hidden information method in Table 49.

Table 49: Repeat of Table 47, macro hidden Information for the SCOR Analysis Levels.

	Complete Matrix Hidden Information	Plan	Source	Make	Deliver	Return	NGO
Level Three	7.53916	4.90689	4.75488	4.95419	5.93073	5.16992	0
Level Two	6.10852	3.90689	3.70044	3.58496	3.70044	3.90689	0
Level One	3.45943	1	1	1	1	1	0
Level Zero	2.58496	0	0	0	0	0	0

A graphical comparison of entropy and hidden information at all levels and for the complete adjacency matrix is shown in Figure 43.

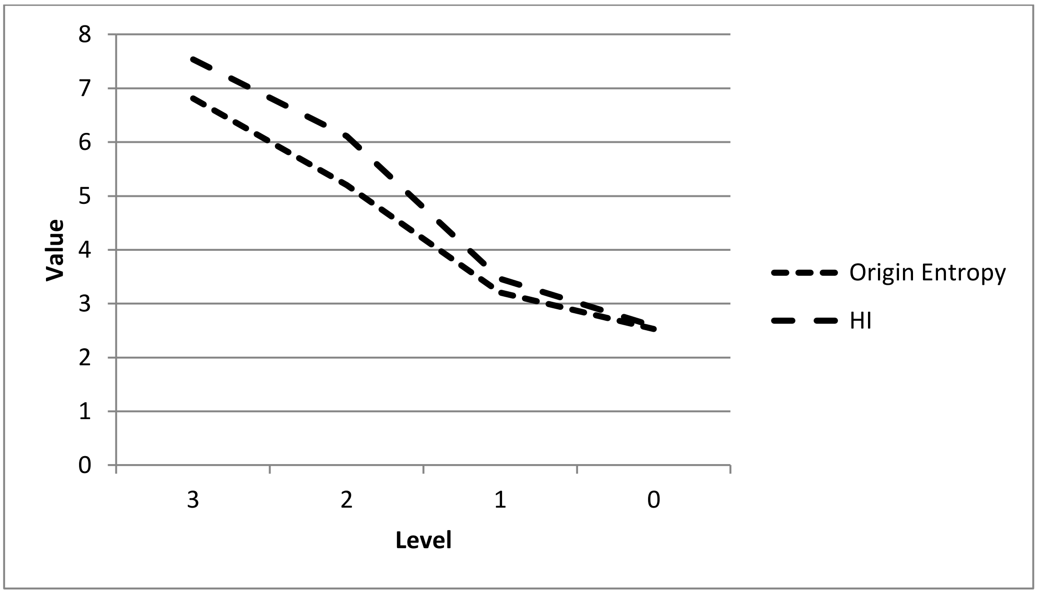


Figure 43: Graphical comparision of entropy and hidden information for the complete adjacency matrix

As can be seen, the two values correlate. The pearson correlation co-efficient for m_k aggregation groupings supports this assertion and is statistically significant at alpha = 5 ($r_{cv} = 0.95$) for entropy $r_{obt} = 0.99607,$ thus, $r(2) = +0.99607, p < .05\ (two - tailed)$.

This correlation continues for each of the plan, source, make, deliver and return subset. Figure 44 to Figure 48 on the following page demonstrate this continuation, again with the x axis representing the level of analysis and y the numeric value.

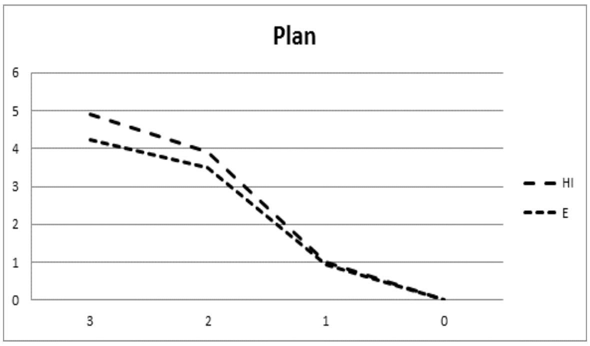


Figure 44: Entropy and Hidden Information for the Plan subset

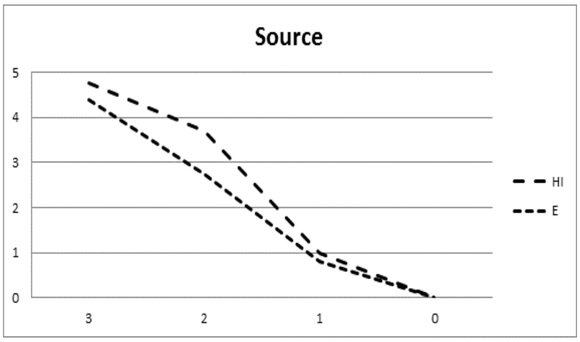


Figure 47: Entropy and Hidden Information for the Source subset

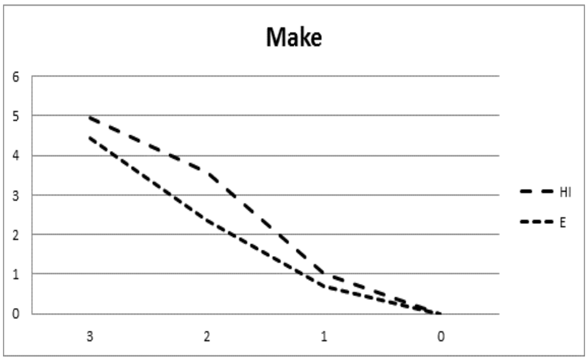


Figure 45: Entropy and Hidden Information for the Make subset

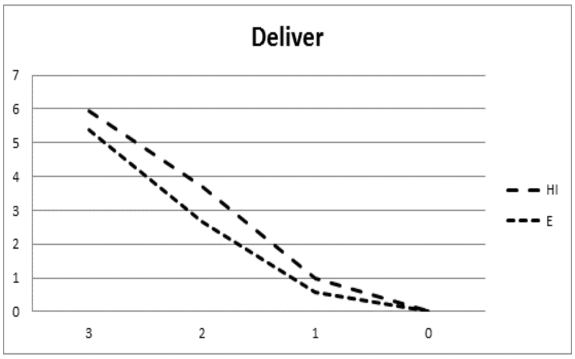


Figure 48: Entropy and Hidden Information for the Deliver subset

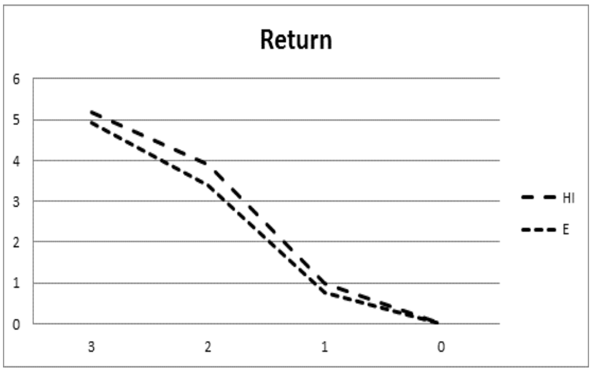


Figure 46: Entropy and Hidden Information for the Return subset

The Pearson correlation coefficients for each sub section are set out in Table 50:

Table 50: Table of subset correlation coefficients.

Subset	Coefficient
Plan	0.99946
Source	0.98780
Make	0.98114
Deliver	0.99113
Return	0.99794

Thus, assuming an alternative hypothesis to be there to be no statistically significant correlation between entropy at the macro level and hidden information, and the null hypothesis to be that a statistically significant correlation exists. Pearson product moment correlation co-efficient for each value is set out in Table 51: $r_{obt} = 0.41$ compares to a two tailed test value for alpha = 0.5 $r_{cv} = \pm 0.4438$. Thus $r(18) = +0.41, p > .05\ (two - tailed)$; showing there to be a statistically significant correlation between entropy and hidden information at the macro level.

Table 51: Pearson product moment correlation.

Subset	Coefficient ($\boldsymbol{r_{obt}}$)	alpha ($\boldsymbol{r_{cv}}$)	
Plan	0.99946	0.95	$r(2) = +0.99946, p > .05\ (two - tailed)$
Source	0.98780	0.95	$r(2) = +0.98780, p > .05\ (two - tailed)$
Make	0.98114	0.95	$r(2) = +0.98114, p > .05\ (two - tailed)$
Deliver	0.99113	0.95	$r(2) = +0.99113, p > .05\ (two - tailed)$
Return	0.99794	0.96	$r(2) = +0.99794, p > .05\ (two - tailed)$

The above describes the relationship between entropy and hidden information at the macro level. To remind the reader, the macro level is the ‘horizontal’ view of the business as described in Figure 9 and Figure 10 on page 66. The reader will notice that a graph for the NGO values has not been included. This is simply because entropy for NGO is a function of 116 K_{out} degree over the possible 1261 links in the matrix. As the observed level transitions up or down the different level is the matrix structure, for all the subsections except NGO, data is aggregated (or disaggregated) and the entropy value changes. For NGO this is not the case; at each level in the matrix, because no aggregation occurs, the values remain the same. By comparision, NGO for hidden information is zero and remains so thoughout the different levels. For hidden information there is only one value for NGO at any level in the matrix and $\log_2 1 = 0$; that is, there is no question to ask over the state of the business if, in asking an average of 2.58496 questions at level zero for the complete matrix, NGO has been determined as the state. Of course, the NGO state is an aggregation of the 116 micro states that make up the NGO group and so, even if NGO has been determined through a macro analysis there remain a number of micro analysis question necessary to determine the actual detailed state the business is in. This analysis is the subject of the next section that describes the relationship between entropy and hidden information at the micro level; the vertical component in Figure 9 and Figure 10.

5.2.3 Micro Level Analysis for the Basic Construct

The micro level analysis looks at the 'vertical' component of information aggregation as described in Figure 9 and Figure 10 on page 66. To do this, the research will compare the proposed hidden information function

$$Q_m = \sum_{1}^{k} \frac{\log_2 a^n}{k} \qquad \text{Equation 74}$$

introduced in chapter two, with the function

$$S = \frac{\sum \frac{\mathrm{p_n}}{\mathrm{p_{m_k}}} \log_2 \frac{\mathrm{p_n}}{\mathrm{p_{m_k}}}}{p_{m_k}} \qquad \text{Equation 75}$$

also introduced and explained in chapter two (Equation 29), which represents the difference between entropy at each level of aggregation. As such

$$S_d = \frac{\sum \frac{\mathrm{p_n}}{\mathrm{p_{m_k}}} \log_2 \frac{\mathrm{p_n}}{\mathrm{p_{m_k}}}}{p_{m_k}} = S_{\ o} - \ S_a \qquad \text{Equation 76}$$

where S_o is the base level entropy and S_a the entropy at aggregation level a.

Comparison of the vertical component follows the same format as the horizontal component; but in this case the values are achieved through the application of the above formulae and the to the changing independent variable *k* for each level of aggregation.

For entropy, $S_d = S_o - S_a$, thus, the values in Table 52 can be derived from Table 45 for example, at level two, $1.6141(S_d)$ = $6.8136\ (S_o) - 5.1995\ (S_a)$. All other values in the following table can be derived using the same approach.

Table 52: Difference micro entropy for the complete matrix.

	Complete Matrix Entropy Difference (S_d)	Plan	Source	Make	Deliver	Return	NGO
Level Three	0	0	0	0	0	0	0
Level Two	1.6141	0.08804	0.31611	0.33982	0.54382	0.32631	0
Level One	3.60433	0.39403	0.68394	0.67039	0.96335	0.89259	0
Level Zero	4.28545	0.50907	0.83708	0.79454	1.08195	1.06271	0

Hidden information, because of its non-additive characteristic, cannot be derived using a subtraction method; instead Equation 74 will be applied to the adjacency matrix to determine the values in Table 53.

Table 53: Difference hidden information for the complete matrix.

	Complete Matrix Hidden Information Difference	Plan	Source	Make	Deliver	Return	NGO
Level Three	2.05314	1.55193	2.63680	2.30373	1.49988	2.57815	6.85798
Level Two	3.11935	2.82305	3.05589	3.00517	2.91933	3.48619	6.85798
Level One	6.56405	6.20137	6.69124	6.45794	6.46905	6.85368	6.85798
Level Zero	7.65240	7.24793	7.90689	7.82018	7.98868	8.09276	6.85798

This vertical/micro analysis highlights and number of points: Firstly, the subtraction method used to determine entropy$(S_d = S_{\ o} - \ S_a)$ demonstrates, for level three, there to be no aggregation. From there entropy increases as aggregation occurs though the levels of analysis. Entropy for each section follows the same pattern; maintaining the additive nature of the concept. Table 53 shows the equivalent values for hidden information. These values, while rising in the same way as entropy, begin with a positive value, unlike the zero values for entropy. To continue to justify hidden information as a valid construct this need some explanation.

Chapter two described the horizontal and vertical construct concept and explained how the construct functions; here we see the manifestation of this construct. For hidden information, the horizontal/macro binary question distance for the complete adjacency matrix at level three is 7.53916 (see Table 49) questions; that is, it will take this number of binary question to understand in what process state the business is in at this level of information aggregation. However, even when this is complete, the vertical/micro analysis in Table 53 tells the questioner that even with the macro information; he or she remains 2.053 binary questions away from understanding the true state of the business. This is because the model aggregates from a deeper set of variables (K_{out} in this research) that are unevenly distributed across the SCOR processes; the value 2.053 provides a measure of the average binary question distance, from the processes, to these underlying values. Put another way, the horizontal/macro questions – for a given level of aggregation – point the questioner to the state of the business as being in one of the processes. Once there, the vertical/micro questions – for the same level of aggregation – indicate to the questioner the average number of questions he or she still needs to ask to get to a specific state within the process set. This may seem complicated, but it is just the same as the example using two dice in chapter two. The macro analysis tells the player how many question would need to be ask to determine, for instance, seven as a score; the micro analysis tells the player the average number of questions required to determine which of the six possible combinations that sum to seven had in fact been scored. This difference is not visible to entropy because there is no 'gap' created by aggregation at this level. Take level zero as a further example. At this level (Table 49 again), the horizontal/macro binary question distance for the complete adjacency matrix is 2.58496 questions; at this level, because of the reduced number of variables caused by the aggregation of data in the hierarchy, it will only take this number of questions to determine the state of the business – the questioner has only a small number of variables from which to choose. However, the micro analysis - Table 53 - shows a value of 7.65240, meaning that while it took fewer questions from the horizontal/macro perspective, there is a greater binary question distance from the vertical/micro perspective to determine the state of the system.

A second observation is the vertical/micro hidden information values for the subsections are similar in value to the whole matrix (See Table 53). How could it be that it will take a similar number of binary questions to find the state of a subsection of the business, as it would to find the state of the whole business? This can explained by comparing values at horizontal/macro and vertical/micro levels. Table 54 takes the level zero data from Table 49 and Table 53 respectively.

Table 54: Hidden information for level zero at horizonal/macro and vertical/micro levels..

	Complete Matrix	Plan	Source	Make	Delivery	Return	NGO
Level Zero micro Hidden Information from (Table 53)	7.65240	7.24793	7.90689	7.82018	7.98868	8.09276	6.85798
Level Zero macro Hidden Information from (Table 49)	2.58496	0	0	0	0	0	0

What this shows is that at level zero in the matrix it takes, on average, 7.65240 additional binary questions to get to the state of the system after asking, on average, 2.58496 macro questions. Whereas, for each of the subsets, there are no macro question; that is, there is only one level zero macro state in the matrix for each subsection; but there will be the maximum number of micro questions for each state. The values ‘trade off’ inversely. Allowing for rounding, the vertical/micro value for the complete matrix should approximate to the average of the values for each of the subsections; in this case the average value of the subsections is 7.7333, a difference of 1.05%. The NGO values follows the same logic except that values remain the same irrespective of the level in the matrix, simply because there is only one NGO category at each level in the matrix; consequently the horizontal/macro value is always zero and the vertical/micro value is always 6.85798 in this case.

A third observation is that a key determinant of the entropy values is the reduction – through aggregation – in the number of variables, from 186 at level three, 69 at level two, 11 at level one to 6 at level zero. Take, for instance, the plan subset as an example: In Table 45 the macro values for matrix levels zero and three are 0.87701 and 0.36794 respectively, with the difference ($S_d = S_o - S_a$) being 0.50901 displayed in Table 52. This difference arises because entropy is calculated as the sum of $log\ (2)$ for 30 k_{out} variables at level three, whereas for level zero, these value are aggregated into just one variable with entropy calculated against this one value. This results in a lower value of entropy, following a general principle that entropy reduces as the number of variables reduce – things get more certain, entropy has the property of extensivity; except in the scenario certainty is created by aggregation of information.

The graphical comparison of hidden information and entropy at the micro level and for the complete matrix is shown in Figure 49.

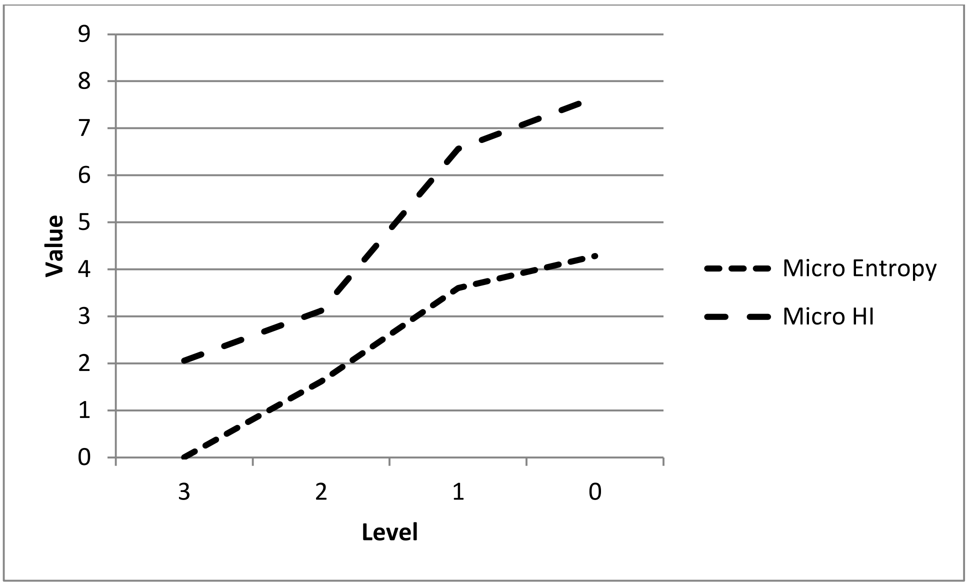


Figure 49: Entropy and hidden information 'micro level' for the complete matrix.

The Pearson correlation co-efficient for the data = 0.98332; thus, assuming an alternative hypothesis to be there is no statistically significant correlation between entropy at the micro level and hidden information, and the null hypothesis to be that a statistically significant correlation exists. Pearson product moment correlation co-efficient: $r_{obt} = 0.98332$ compares to a two tailed test value for alpha = 0.5 $r_{cv} = \pm 0.9500$. Thus $r(2) = +0.98332, p > .05\ (two - tailed),$ demonstrating statistical significance between the two values.

Comparing the graphs for horizontal/macro entropy (Figure 43) with the graph for the vertical/micro entropy analysis (Figure 49) shows the slopes of each to be inverse. For horizontal/macro analysis, as data is aggregated, values reduce; whereas, for vertical/micro analysis, as data is aggregated, values increase. Intuitively, this makes sense as the trade-off between the two. From a vertical/micro perspective, the more aggregation, the greater the binary question distance; from a horizontal/macro perspective, the more aggregation, the lesser the number of binary questions. Conversely, a reduction in aggregation reduces the number of vertical/micro binary question and increases the number of horizontal/macro questions.

This section has analysed the complete adjacency matrix at the basic, macro and micro levels where all the processes are used. The complete matrix shows a skewed distribution for the K_{out} links, The macro analysis demonstrates how, simply by building the K_{out} links into a process structure, entropy and hidden information reduce from their consistent and equiprobable value. Furthermore, this trend continues as the aggregation increases, both for entropy and hidden information. Also, the trend is consistent for each component part of the macro analysis; plan, source, make, deliver and return all show the same trend. The analysis also demonstrates the difference and correlation in entropy values and hidden information for the complete matrix and its components. The micro analysis firstly demonstrates the manifestation of question distance as additional necessary questions and how entropy at the origin process level does not 'see' this distance. In fact, this analysis demonstrates how entropy simply reduces with the number of variables. Finally, the micro analysis allows a graphical comparison with the macro analysis. This shows the opposing graph slopes which demonstrates the trade-off position of the two dimensions of entropy and hidden information that have been the subject of the research so far. The analysis continues to question the validity of entropy as a number of questions, and demonstrates this through the simple additive nature of the value. It also shows how hidden information does not follow this additive structure but does appear to provide values that are a true reflection of the number of questions required both horizontally and vertically. The analysis covers only the complete SCOR process model. To validate these initial results, further testing is required against the defined scenarios. The next section will analyse the validity of these present findings against these predetermined scenarios.

5.3 Comparison of Hidden Information and Entropy for the Multi-organisational Construct

As explained above, this scenario represents the prime contractor role in organisations, where plan and source are completed twice: once for the role of prime contractor, planning and outsourcing various commodities, and once for the internally made in components. To represent this scenario plan and source are repeated in the business model. To model this scenario correctly probability p is now taken across the wider set of possible states:

$$p_n = \frac{k_{nout}}{\sum_1^n k} \qquad \text{Equation 77}$$

where k_{nout} = the out degree for each of the plan, source, make, deliver and return processes, plus the additional plan and source processes for the prime contractor role in this scenario. Similarly, for hidden information, $n \in N$ equates to the same plan, source, make, deliver and return processes plus the additional plan and source processes. For clarity the horizontal/macro and vertical/micro analyses will again be broken down into two separate sections beginning with the horizontal/macro analysis. A pictorial representation of the scenario is shown in Figure 50.

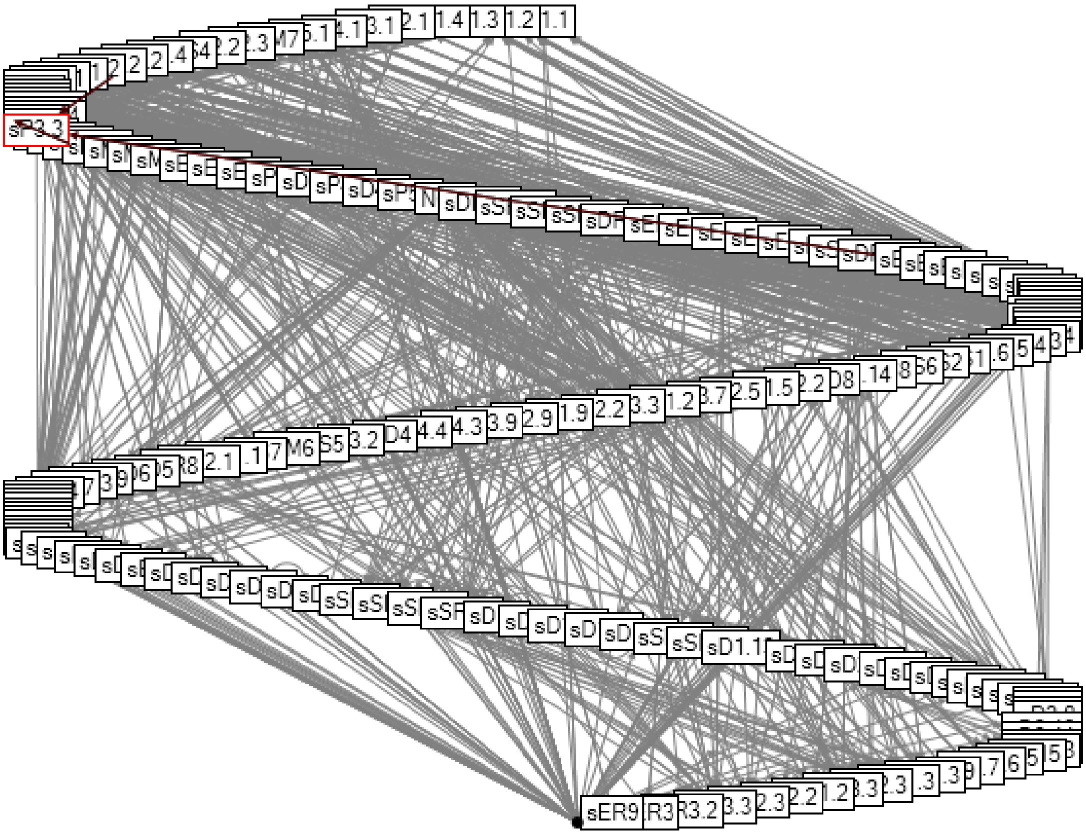


Figure 50: Superset network diagram.

5.3.1 Macro Level Analysis of Multi Organisation Construct

The horizontal/macro level analysis for this scenario consists of eight sub-elements forming the overall entropy and hidden information in the business model. Table 55 shows entropy for this scenario at all levels of aggregation and for each of the eight SCOR elements. The table also includes values in brackets; these values represent the number of variables for each element at that given level of aggregation in the process model. At level one the data is aggregated into two values, both have been provided. At level zero all the variables are aggregated to one variable which accounts for there being only one value in the table.

Table 55: Superset macro entropy for the SCOR Model.

	Overall Entropy	Plan	Source	Plan	Source	Make	Deliver	Return	NGO
Level Three	7.24339	0.70494 (30)	1.04278 (27)	0.70494 (30)	1.04278 (27)	0.99860 (31)	1.24058 (61)	1.23978 (36)	0.26897 (1)
Level Two	5.70375	0.66378 (15)	0.80163 (13)	0.66378 (15)	0.80163 (13)	0.73937 (12)	0.82573 (13)	0.99085 (15)	0.26897 (1)
Level One	3.67147	0.23684 0.16752	0..34941 0.17162	0.23684 0.16752	0.34941 0.17162	0.35256 0.13463	0.38545 0.12024	0.37723 0.18162	0.26897
Level Zero	2.94737	0.31659	0.40421	0.31659	0.40421	0.39248	0.41522	0.42909	0.26897

Hidden information, for the same scenario is set out in Table 56.

Table 56: Superset macro hidden information for the SCOR model.

	Overall Hidden Information	Plan (1)	Source (1)	Plan	Source	Make	Deliver	Return	NGO
Level Three	7.92481	4.90689	4.75488	4.90689	4.75488	4.95419	5.93073	5.16992	0
Level Two	6.59991	3.90689	3.70044	3.90689	3.70044	3.58496	3.70044	3.90689	0
Level One	3.90689	1	1	1	1	1	1	1	0
Level Zero	3.0000	0	0	0	0	0	0	0	0

Comparing the values in the two tables, the issue with the use of entropy as an understanding of ‘mixed up ness’ continues to manifest. Take, for example, the results at level zero – maximum aggregation: For entropy, the overall value is 2.94737 or $2^{2.94737} = 7.71342$ variables when in fact there are eight variables, one variable for each sub element. Thus, again, the value of entropy is less than the actual value required to be a valid measure of questions; whereas the value (3) for hidden information demonstrates this as the number of binary questions for eight variables $2^3 = 8$. Entropy is additive, and in this scenario the value for entropy is the sum of eight variables, ranging in value from 0.26897 to 0.42909. These values indicate a level of mixed up ness but, as before, they do not provide a value for the number of binary questions required. At level one aggregation there are two entropy variables for each of the elements; one binary question would be required. The entropy values $(0 < x > 1)$ do not reflect question values. Hidden information does reflect the required value, one in each case. The same holds true for all values in the entropy table at all level of aggregation and all comparisons with hidden information values.

Table 57 compares entropy and hidden information values for the basic analysis with this multi organisation scenario.

Table 57: Comparison value for Origin and the Multi-organisation scenario.

	Origin Entropy	Scenario Entropy	Origin Hidden Information	Scenario Hidden Information
Level Three	6.81355	7.24339	7.5391	7.92481
Level Two	5.19945	5.70375	6.10852	6.59991
Level One	3.20922	3.67147	3.45943	3.90689
Level Zero	2.5281	2.94737	2.58496	3.0000

The various differences in the values in Table 57 are shown in Table 58.

Table 58: Difference in entropy and hidden information values.

	Origin Entropy vs Scenario Entropy	HI Origin vs HI Scenario	Origin Entropy vs HI Origin	Scenario Entropy vs HI Scenario
Level Three	0.42984	0.38571	0.72555	0.68142
Level Two	0.5043	0.49139	0.90907	0.89616
Level One	0.46225	0.44746	0.25021	0.23542
Level Zero	0.41927	0.41504	0.05686	0.05263

For the basic analysis and this scenario, as information is aggregated, the number of variables reduces. Table 59 shows this reduction in variables over each layer of aggregation with column three showing the additional number of variables in this scenario.

Table 59: Reduction in variables in the matrix.

	Basic Analysis Volumes	Percentage Reduction	Scenario Volumes	Percentage Reduction
Level Three	186		186+58= 245	
Level Two	69	62.9032	69+28= 97	60.4082
Level One	11	84.0580	11+4= 15	84.5361
Level Zero	6	45.4545	6+2=8	46.6667

The values in Tables 58 - 60 demonstrate an alignment with a number of properties referred to in earlier sections. Firstly the additive property of entropy continues, whereas hidden information does not display the same property; it is this non-additive property that contributes to the ability of hidden information to create of binary question values at all levels of aggregation and for all process scopes. Secondly, comparing hidden information for this scenario (Table 56) with the same values for the basic analysis (Table 46) demonstrates consistency in the hidden information approach. Based on $\log_2$ for n variables, the method provides clarity on the question distance for macro analysis. The method omits the application of probability and thus more complex results given by the entropy method.

Comparing entropy for the basic analysis and this scenario (Table 55 and Table 45) with hidden information for the same scenarios (Table 56 and Table 46) shows consistency of common extensive and monotonic characteristics, and the additive/non-additive difference between the two methods; that is, entropy is additive and hidden information is non-additive. The extensive characteristic in both cases relates to the number of variables; both entropy and hidden information change in line with a change in the volume of data elements. For the cases in this research this is due to data aggregation, but in the general case the same principle is applicable where volumes, aggregation or otherwise, reduce. Being additive entropy demonstrates a counter intuitive characteristic present when elements of the model are added – as in this scenario - to the organisational construct. Take for example, the Plan element of the SCOR model. For the basic analysis at level three entropy = 0.87701; the same value for this multi – organisation scenario = 0.70494. The reduction implies a reduction in complexity: entropy is reducing, therefore certainty must be increasing; but this is not the case, uncertainty has increased with the increase in possible states for the organisation. In reality, the reason for this apparent reduction in complexity is this increase in the total number of variables, which reduces the probability of the use of any one specific link in the organisation structure. The additive nature reduces the value of each element thus allowing for an increase in the number of elements necessary for the construction of the scenario. By comparison, hidden information for each element is unchanged, demonstrating the ‘distance’ to be unchanged by the number of elements. Throughout, entropy and hidden information for the scenario correlate.

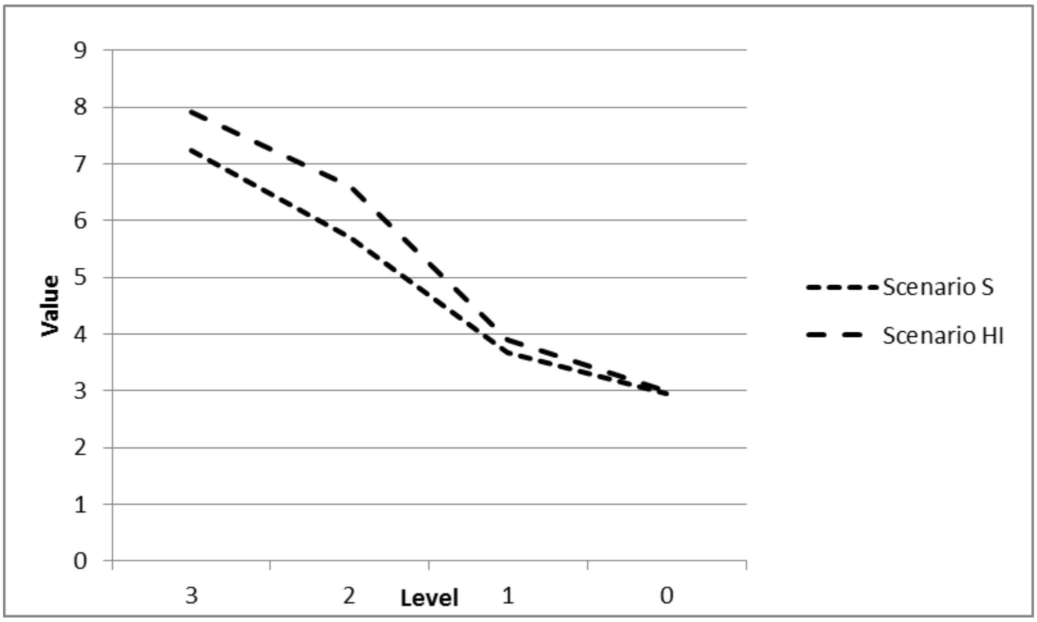


Figure 51: Scenario Comparison Entropy and Hidden Information

The Pearson correlation co-efficient for the scenario data $r_{obt} = 0.99598$ is statistically significant, thus, $r(2) = +0.99598, p > .05\ (two - tailed)$.

This section set out to compare macro level entropy and hidden information for the superset scenario which represents the prime contractor role in the supply chain. The key findings are: both measures display extensive and monotonic properties, entropy continues to demonstrate its additive property and hidden information its non-additive property. The exclusion of probability in the hidden information method continues to demonstrate the validity of the approach; that is, $\log_2 N$ provides a more quantifiable value in that a unit of measure is applied to the measure. The extensive property of both entropy and hidden information are a function of the volume of variables at each level of aggregation. The multi-organisation scenario can be considered statistically independent of the basic analysis. Finally, entropy reduces for the same structural element if the number of elements increases, creating an issue due to the additive nature of the measure when applied to a structural model. Additionally, this problem emphasises the need for a combined adjacent and orthogonal measure of complexity for the same structure. This will be further analysed in the next section.

5.3.2 Micro Level Analysis for the Multi Organisation Construct

As was the case with the previous analysis, this section sets out the micro level analysis; that is, the creation of values orthogonal to the values in the previous section. Again this analysis will compare values from the proposed hidden information function $Q_m = \sum_1^k \frac{\log_2 a^n}{k}$, from Equation 74, with the function $S = \frac{\sum \frac{\mathrm{p_n}}{\mathrm{p_{m_k}}} \log_2 \frac{\mathrm{p_n}}{\mathrm{p_{m_k}}}}{p_{m_k}}$, from Equation 75, but in this case the processes will be inclusive of the additional plan and source elements identified for this scenario.

Table 60: Entropy values for the micro-analysis of the multi-organisational scenario.

	Overall Entropy	Plan	Source	Plan	Source	Make	Deliver	Return	NGO
Level Three	0	0	0	0	0	0	0	0	0
Level Two	1.53964	0.04116	0.24115	0.04116	0.24115	0.25923	0.41485	0.24893	0
Level One	3.57192	0.30058	0.52175	0.30058	0.52175	0.51141	0.73489	0.68093	0
Level Zero	4.29602	0.38835	0.63857	0.38835	0.63857	0.60612	0.82536	0.81069	0

For entropy, being additive, $S_d = S_o - S_a$, thus, the values in Table 60 can be derived from Table 55. For example, at level two, overall entropy $S_d = 1.53964$ is derived as $1.53964(S_d) = 7.24339(S_o) - 5.70375(S_a)$. All other values in Table 60 are derived similarly. Hidden information values, because of their non-additive nature, cannot be derived using the same approach; consequently the formula developed in chapter two and described above has to be applied to the adjacency matrix in order to determine the values in Table 61.

Table 61: Hidden information for the micro analysis of the multi-organisational scenario.

	Overall Hidden Information	Plan	Source	Plan	Source	Make	Deliver	Return	NGO
Level Three	2.06250	1.55193	2.63680	1.55193	2.63680	2.30373	1.49988	2.57815	6.85798
Level Two	3.06503	2.82305	3.05589	2.82305	3.05589	3.00517	2.91933	3.48619	6.85798
Level One	6.53265	6.20137	6.69124	6.20137	6.69124	6.45794	6.46905	6.85368	6.85798
Level Zero	7.63365	7.24793	7.90689	7.24793	7.90689	7.82018	7.98868	8.09276	6.85798

Two dimensions are available for analysis. Firstly there is the comparison of entropy and hidden information in this scenario; secondly there is the comparison of entropy and hidden information values in this scenario, with those in the earlier basic analysis. Taking the former of the two, comparisons within this scenario highlight a number of analytical points: Zero values at entropy level three compared with values ranging from 1.49988 to 2.63680 for hidden information respectively recognise the embedded structural variation described in the analysis of the basic construct; that is, there is no visible vertical aggregation of data from an entropy perspective, there is no entropy. By comparison, given the variations in the k_{out} degree, there remains a value for hidden information as an average number of additional questions required to identify a specific state. These level three hidden information values again demonstrate the notion that: a) the overall value is the average of the individual values for each of the sub-sets and, b) hidden information remains when the observer is at the most detailed level of analysis available due to the structure of the matrix, which is determined by the architecture of the business process. Later in this chapter we will discuss how these values describe the in-effectivity of functional silos in a business.

Comparing the values in this scenario with those in the basic analysis we see the value for overall hidden information is 2.0625; a small increase over the same value for the basic construct (2.05314). The increase is due to the increase in the number of variables - from 186 to 243 - and the impact this has on the weighted sum over the additional processes. That said, the value does remains consistent with the values determined in the analysis of the basic construct. For hidden information, the horizontal/macro binary question distance for this scenario is 7.92481(Table 56) questions (it will take this number of binary question to understand in what process state the business is in). And even when this is known, the vertical/micro complexity in this scenario tells the questioner that even with this information, he or she remains 2.06250 binary questions away from understanding the state of the business. This is consistent with the basic analysis above and is explained because the model aggregates a deeper set of variables (K_{out} in this research) that are unevenly distributed across the SCOR processes; the value 2.06250 is providing a view of the binary question distance, from these values to the processes. This difference is not visible to entropy because there is no ‘gap’ created by aggregation at this level. Recall, this was explained theoretically in chapter two and practically earlier in this chapter. If the same values are considered for level zero – from Table 56 - the horizontal binary question distance for the complete adjacency matrix is 3.0000 (there are eight variables) questions; that is, at this level, because of the reduced number of variables created through the aggregation of data in the hierarchy, for an increased number of level zero processes that recognise the greater business scope in this scenario, it will take this number of questions to determine the state of the business. At this level, the micro analysis (Table 56) gives a value of 7.63365, which means that while it took fewer questions from the horizontal/macro perspective, this leads to a greater binary question distance from the vertical/micro perspective. All of which is consistent with the analysis of the basic model construction.

An important observation from the analysis of the basic matrix was that hidden information values for the subsets are similar to the value for the whole matrix; i.e. it would take a similar number of binary questions to identify the state of a subset, as it would the whole matrix. For this scenario the values, taken from Table 61 and Table 56 are provided in Table 62.

Table 62: Hidden information for level zero at macro and micro level.

	Overall Hidden Inf,	Plan	Source	Plan	Source	Make	Deliver	Return	NGO
Level Zero Micro HI	7.63365	7.24793	7.90689	7.24793	7.90689	7.82018	7.98868	8.09276	6.85798
Level Zero Macro HI	3.0000	0	0	0	0	0	0	0	0

The earlier analysis of the basic construct resulted in an average 7.65240 additional binary questions to get to the state of the system after asking, on average, 2.58496 macro questions. In both scenarios, for each of the macro subsets, there are no additional questions; there is only one state for each subset at the macro level. Allowing for rounding, the micro value for the complete matrix should approximate to the average of the values for the subsets. In this scenario the results show, on average, 7.63365 additional binary questions to get to the state of the systems after asking, on average, 3.0000 macro questions, matching the complete matrix value. The NGO values follows the same logic with the result that values remain the same irrespective of the level in the matrix, simply because there is only one NGO category at each level in the matrix; consequently the horizontal/macro value is always zero and the vertical/micro value is always 6.85798 in this case.

The key determinant of the entropy values is the reduction in the number of variables. In the basic analysis the process of aggregation reduced the number of variables from 186 at level three to 69 at level two and then 11 and 6 at levels one and zero respectively. In this multi organisation scenario those values have increased to 245, 97, 15 and 8 respectively. Entropy continues to demonstrate a micro level increase as aggregation increases, following ($S_d = S_o - S_a$). In this scenario the actual values are generally lower for the subsets due to the increase in the number of variables and the additive nature of entropy. The k_{out} degree probability for each value and grouped sum decreases due to the increase in the number of variables which reduces the probability values.

Graphically the values for hidden information and entropy are shown in Figure 52.

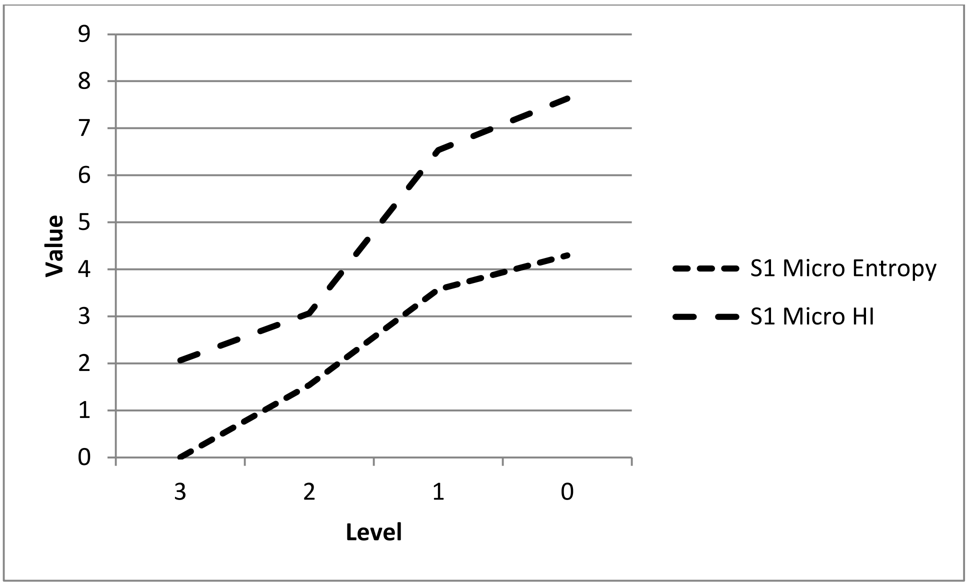


Figure 52: Entropy and hidden information 'micro level' for the multi organization scenario.

The Pearson correlation co-efficient for the data = 0.98484; thus, assuming an alternative hypothesis to be there is no statistically significant correlation between entropy at the micro level and hidden information, and the null hypothesis to be that a statistically significant correlation exists. Pearson product moment correlation co-efficient: $r_{obt} = 0.98484$ compares to a two tailed test value for alpha = 0.5 $r_{cv} = \pm 0.9500$. Thus $r(2) = +0.98484, p > .05\ (two - tailed)$, demonstrating statistical significance between the two values.

Comparing the graphs for horizontal/macro entropy (Figure 51) and those for the vertical/micro entropy analysis (Figure 52) shows the slopes of both hidden information and entropy to be inverse. For horizontal/macro analysis, as data is aggregated, values reduce; whereas, for vertical/micro analysis, as data is aggregated, values increase. Intuitively, this again makes sense: From a vertical/micro perspective, the more aggregation, the greater the binary question distance. From a horizontal/macro perspective, the more aggregation, the lesser the number of binary questions. Conversely, a reduction in aggregation reduces the number of vertical/micro binary question and increases the number of horizontal/macro questions.

This section has analysed the adjacency matrix for the multi organisation scenario; at both the macro and micro levels. The macro analysis continues to demonstrate how, simply by constructing a process structure and attaching K_{out} links, entropy and hidden information reduce from their consistent, equiprobable values. The analysis also demonstrates the correlation and difference in entropy and hidden information values for the scenario matrix. The micro analysis continues to demonstrate and validate the concept of question distance as average additional necessary questions to identify a given state structure; and how entropy at the non-aggregated process level does not ‘see’ this distance. The analysis continues to demonstrate how entropy simply changes with the number of variables with figures that are unquantified. Finally, the graphical analysis shows the slope of the graphs for micro analysis to be inverse to those for macro analysis, demonstrating again the two dimensions of entropy and hidden information that have been the subject of the research so far. The multi-organisation scenario shows results that are consistent, with respect to entropy and hidden information, to those of the basic analysis. And it shows how the hidden information measure continues to provide definitive and measurable information. Through the use of this scenario, the analysis continues to show entropy to be an indicator of ‘mixed up ness’, but not to be a measurable value of questions or question distance; further, it shows how hidden information does not follow the additive restriction and does appear to provide values that are a true reflection of the number of questions required both horizontally and vertically. The analyses thus far have covered basic and multi organisation scenarios. To further validate these initial results, further testing is required against the subcontractor role, represented in the minimalist scenario in the next section.

5.4 Comparison of Hidden Information and Entropy for Minimalistic Construct

In this scenario the subcontractor role, where plan and source activities are completed by a prime contractor and the role of the subcontractor is to complete the make, deliver and return processes, will be studied. To represent this scenario, plan and source activities have been excluded and are represented within the analysis as sections with no processes. As previous, probability p is taken across all possible business states, thus:

$$p_n = \frac{k_{nout}}{\sum_1^n k} \qquad \text{Equation 78}$$

where k = the out degree for the make, deliver and return processes. Similarly, for hidden information, $n \in N$ equates to the same make, deliver and return processes. For clarity the macro and micro analysis will be broken down – as previously - into two separate sections, beginning with the macro analysis. Pictorially, the scenario is shown in Figure 53.

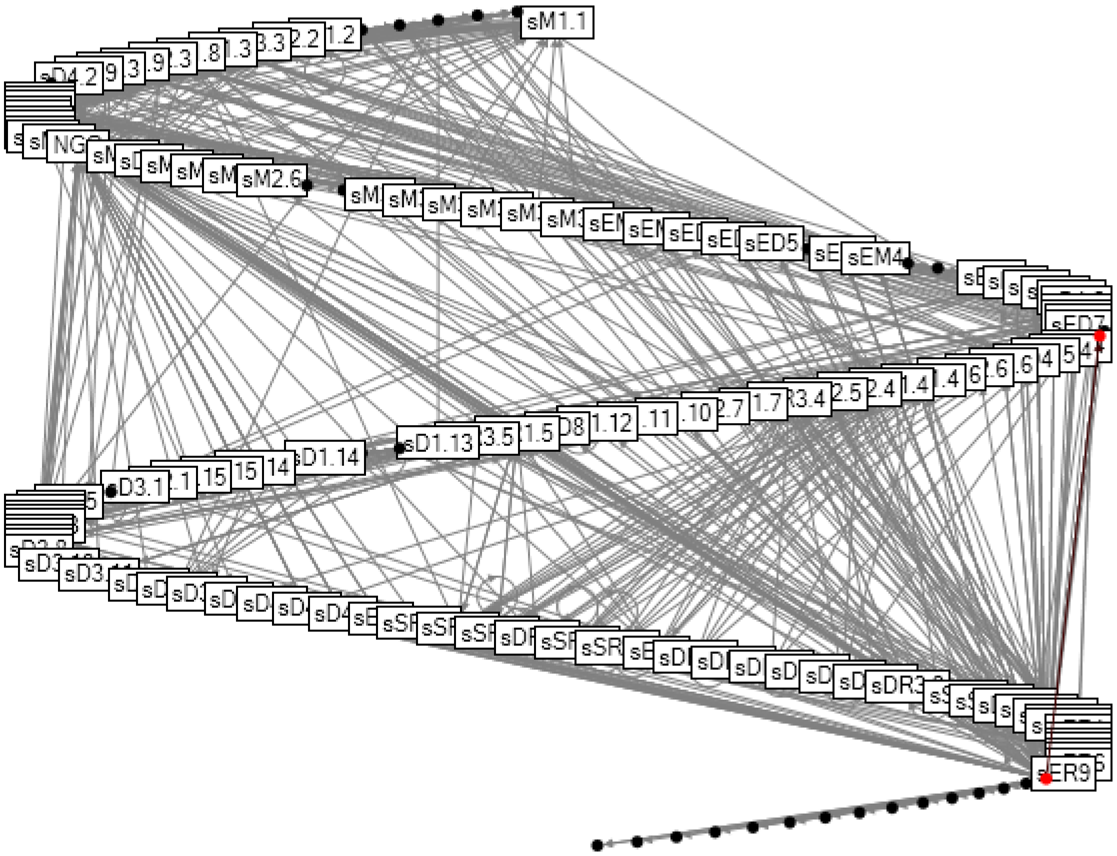


Figure 53: Subset network.

From the diagram it is worth noting how limiting the range of processes to those within the subcontractor scenario begins to affect the coherency of the structure as out- bound links begin to point towards processes that are excluded from the scenario. Specifically, this issue is not for this research; however, the removal of a set of processes can remove, for in scope processes, link destinations. This impacts on coherency and it should be noted that, when designing such solutions, effort in business function process map integration is required to avoid such incoherency.

5.4.1 Macro Level Analysis of Minimalistic Construct

The horizontal/macro level analysis for this scenario consists of four sub-elements forming the overall entropy and hidden information in the business model. Table 63 shows entropy for this scenario at all levels of aggregation and for each of the five SCOR elements. The table also includes values in brackets; these values represent the number of variables for each element at each level of aggregation in the process model. At level one the data is aggregated into two values, both have been provided. At level zero all the variables are aggregated to one variable which accounts for there being only one value in the table.

Table 63: Entropy for the subcontractor scenario.

	Overall Entropy	Plan	Source	Make	Deliver	Return	NGO
Level Three	6.20162	0 (30)	0 (27)	1.65827 (31)	2.08868 (61)	2.06686 (36)	0.38781 (1)
Level Two	4.44588	0 (15)	0 (13)	1.16516 (12)	1.29955 (13)	1.59336 (15)	0.38781 (1)
Level One	2.53564	0	0	0.47422 0.21125	0.50048 0.19029	0.49445 0.27715	0.38781
Level Zero	1.93659	0	0	0.50532	0.51868	0.52478	0.38781

Macro level analysis, set out in chapter two, proposed a comparison of traditional entropy with the proposed form $HI = Log_2N$, hidden information. Hidden information, for the same scenario is shown in the Table 64.

Table 64: Hidden Information for the subcontractor scenario.

	Overall Hidden Information	Plan	Source	Make	Deliver	Return	NGO
Level Three	7.01123	0	0	4.95419	5.93073	5.16992	0
Level Two	5.35755	0	0	3.58496	3.70044	3.90689	0
Level One	2.80735	0	0	1	1	1	0
Level Zero	2.00000	0	0	0	0	0	0

Comparing the values in the two tables, the issue previously identified with the use of entropy as an understanding of ‘mixed up ness’ continues to manifest. Take, in this scenario, the results at level zero – maximum aggregation: For entropy, the overall value is 1.93659 or $2^{1.93659} = 3.828$ variables when in fact there are four variables, one variable for each sub element. Thus, again, the value of entropy is less than the actual value required to be a valid measure of questions; whereas the value (2) for hidden information demonstrates this as the number of binary questions for four variables $2^2 = 4$. Entropy is additive, and in this scenario the value for entropy is the sum of four variables, ranging in value from 0.38781 to 0.52478. These values indicate a level of mixed up ness but, as before, they do not provide a value for the number of binary questions required. At level one aggregation there are two entropy variables for each of the elements; one binary question would be required. The entropy values $(0 < x > 1)$ do not reflect question values. Again, hidden information does reflect the required value, one in each case. The same holds true for all values in the entropy table at all level of aggregation and all comparisons with hidden information values.

Table 65 compares entropy and hidden information values for the basic analysis with this subcontractor scenario.

Table 65: Comparison value for Origin and the subcontractor scenario.

	Origin Entropy	Scenario Entropy	Origin Hidden Information	Scenario Hidden Information
Level Three	6.81355	6.20162	7.5391	7.01123
Level Two	5.19945	4.44588	6.10852	5.35755
Level One	3.20922	2.53564	3.45943	2.80735
Level Zero	2.5281	1.93659	2.58496	2.00000

The various differences in the values in Table 65 are shown in Table 66.

Table 66: Difference in entropy and hidden information values.

	Origin Entropy vs Scenario Entropy	HI Origin vs HI Scenario	Origin Entropy vs HI Origin	Scenario Entropy vs HI Scenario
Level Three	-0.61193	-0.52787	0.72555	0.80961
Level Two	-0.75357	-0.75097	0.90907	0.91167
Level One	-0.67358	-0.65208	0.25021	0.27171
Level Zero	-0.59151	-0.58496	0.05686	0.06341

For the basic analysis and this scenario, as information is aggregated, the number of variables reduces. Table 67 shows this reduction in variables over each layer of aggregation with column three showing the reduction in the number of variables in this scenario.

Table 67: Reduction in variables in the matrix.

	Basic Analysis Volumes	Percentage Reduction	Scenario Volumes	Percentage Reduction
Level Three	186		186-57= 129	
Level Two	69	62.9032%	69-28= 41	68.2171
Level One	11	84.0580%	11-4= 7	82.9268
Level Zero	6	45.4545%	6-2=4	42.8571

The values in Tables 66 - 68 demonstrate an alignment with a number of properties referred to in earlier sections. Firstly, the additive property of entropy continues, whereas hidden information does not display this property; it is this non-additive property that contributes to the ability of hidden information to create binary question values at all levels of aggregation and for all process scopes. Secondly, comparing hidden information for this scenario (Table 64) with the same values for the basic analysis (Table 46) demonstrates consistency in the hidden information approach. Based on $\log_2$ for n variables, the method provides clarity on the question distance for macro analysis. The method omits the application of probability and thus more complex results given by the entropy method.

Comparing entropy for the basic analysis and this scenario (Table 63 and Table 45) with hidden information for the same scenarios (Table 64 and Table 46) shows consistency of common extensive and monotonic characteristics, and the additive/non-additive difference between the two methods; that is, entropy is additive and hidden information is non-additive. In both cases the extensive characteristic relates to the number of variables; both entropy and hidden information change in line with a change in the volume of data elements. For this research this is due to data aggregation, but in the general case the same principle is applicable where volumes, aggregation or otherwise, reduce.

Being additive entropy continues to demonstrate the counter intuitive characteristic present when elements of the model are added or removed, – removed in this scenario - to the organisational construct. Take for example, the Make element of the SCOR model. For the basic analysis at level three entropy = 1.23904; the same value for this multi – organisation scenario = 1.65827. The increase implies an increase in complexity: entropy is increasing, therefore certainty must be decreasing; but this is not the case, uncertainty has decreased with the decrease in possible states for the organisation. In reality, the reason for this apparent increase in complexity is this decrease in the total number of variables, which increases the probability of the use of any one specific link in the organisation structure. The additive nature increases the value of each element thus allowing for a decrease in the number of elements necessary for the construction of the scenario. By comparison, hidden information for each element is unchanged, demonstrating the ‘distance’ to be unchanged by the number of elements. Throughout, entropy and hidden information for the scenario correlate:

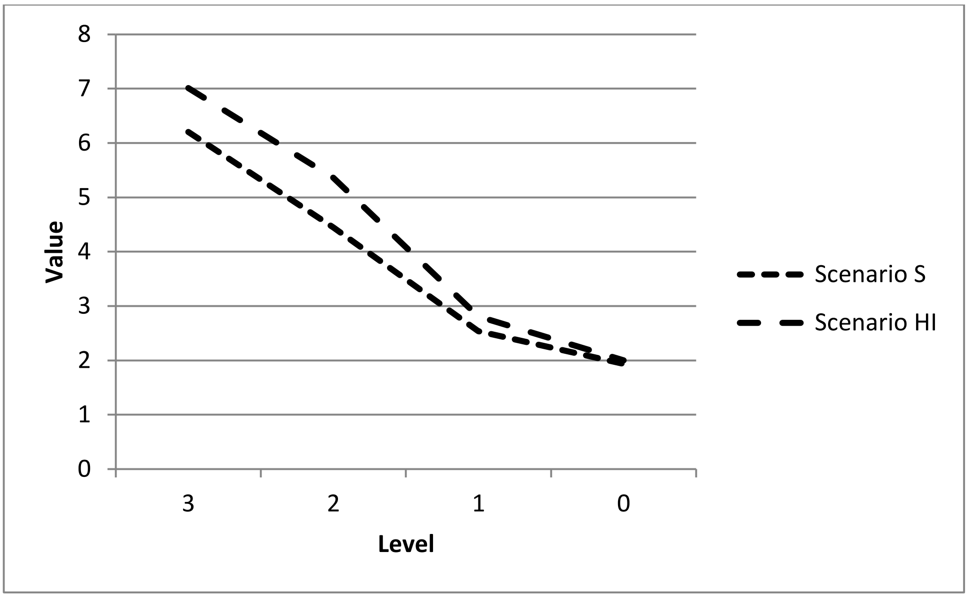


Figure 54: Scenario Comparison Entropy and Hidden Information

As can be seen, the two values correlate. The Pearson correlation co-efficient for m_k aggregation groupings supports this assertion and is statistically significant at alpha = 5 ($r_{cv} = 0.95$) for entropy $r_{obt} = 0.99653,$ thus, $r(2) = +0.99653, p < .05\ (two - tailed)$.

This section set out to compare macro level entropy and hidden information for the subset scenario which represents the sub-contractor role in the supply chain. The key findings are: both measures continue display extensive and monotonic properties, entropy continues to demonstrate its additive property and hidden information its non-additive property. The exclusion of probability in the hidden information method continues to demonstrate the validity of the approach in that $\log_2 N$ provides a more quantifiable value in that a unit of measure is applied to the measure. The extensive property of both entropy and hidden information are a function of the volume of variables at each level of aggregation. Finally, entropy increases for the same structural element if the number of elements decreases, creating an issue due to the additive nature of the measure when applied to a structural model. Additionally, this problem emphasises the need for a combined adjacent and orthogonal measure of complexity for the same structure. This will be further analysed in the next section.

5.4.2 Micro Level Analysis of Minimalistic Construct

As was the case with the basic and prime contractor analysis, this section sets out the micro level analysis; that is, the creation of values orthogonal to the values in the previous section. Again this analysis will compare values from the proposed hidden information function $Q_m = \sum_1^k \frac{\log_2 a^n}{k}$, from Equation 74, with the function $S = \frac{\sum \frac{\mathrm{p_n}}{\mathrm{p_{m_k}}} \log_2 \frac{\mathrm{p_n}}{\mathrm{p_{m_k}}}}{p_{m_k}}$, from Equation 75, but in this case the processes will be exclude the plan and source elements identified for this scenario.

Table 68: Entropy values for the micro-analysis of the sub-contractor scenario.

	Overall Entropy	Plan	Source	Make	Deliver	Return	NGO
Level Three	0	0 (30)	0 (27)	0 (31)	0 (61)	0 (36)	0 (1)
Level Two	1.75574	0 (15)	0 (13)	0.49311 (12)	0.78913 (13)	0.4735 (15)	0 (1)
Level One	3.66598	0	0	0.9728	1.5882	1.29526	0
Level Zero	4.26503	0	0	1.15295	1.57000	1.54208	0

For entropy, being additive, $S_d = S_o - S_a$, the values in Table 68 can be derived from Table 63. For example, at level two, overall entropy $S_d = 1.75574$ is derived from Table 63 as $1.75574(S_d) = 6.20162(S_o) - 4.44588(S_a)$. All other values in Table 68 are similarly derived. Hidden information values, because of their non-additive nature, cannot be derived using the same approach; consequently the formula developed in chapter two and described above has to be applied to the adjacency matrix in order to determine the values in Table 69.

Table 69: Hidden Information for the subcontractor scenario.

	Overall Hidden Information	Plan	Source	Make	Deliver	Return	NGO
Level Three	2.03550	0	0	2.30373	1.49988	2.57815	6.85798
Level Two	3.22479	0	0	3.00517	2.91933	3.48619	6.85798
Level One	6.63133	0	0	6.45794	6.46905	6.85368	6.85798
Level Zero	7.68990	0	0	7.82018	7.98868	8.09276	6.85798

Again, two dimensions are available for analysis. Firstly there is the comparison of entropy and hidden information in this scenario; secondly there is the comparison of entropy and hidden information values for this scenario, with those in the earlier basic analysis. Taking the former of the two, comparisons within this scenario highlight a number of points: Zero values at entropy level three compared with values ranging from 1.49988 to 2.57815 for hidden information respectively again recognise the embedded structural variation described in the analysis of the basic construct; that is, there is no visible vertical aggregation of data from an entropy perspective, there is no entropy. By comparison, given the variations in the k_{out} degree, there remains a value for hidden information as an average number of additional questions required to identify a specific state. These level three hidden information values again demonstrate the notion that: a) the overall value is the average of the individual values for each of the sub-sets and, b) hidden information remains when the observer is at the most detailed level of analysis available within the structure of the matrix, which is determined by the architecture of the business process. Later in this chapter we will discuss how these values describe the in-effectivity of functional silos in a business.

Comparing the values in this scenario with those in the basic analysis we see the value for overall hidden information is 2.0355; a small decrease in the same value for the basic construct (2.05314). The decrease is due to a decrease in the number of variables - from 186 to 129 – which has an impact on the weighted sum over the additional processes; that said, the value remains consistent with the basic analysis in as much as it aligns with the values described in chapter two: For hidden information, the horizontal/macro binary question distance for this scenario is 7.01123 (Table 64) questions (it will take this number of binary question to understand in what process state the business is in). And even when this is known, the vertical/micro complexity in this scenario tells the questioner that even with this information, he or she remains 2.03550 binary questions away from understanding the state of the business. This is consistent with the basic analysis above and can be explained because the model aggregates a deeper set of variables (K_{out} in this research) that are unevenly distributed across the SCOR processes; the value 2.03550 is providing a view of the binary question distance, from these values to the processes. This difference is not visible to entropy because there is no 'gap' created by aggregation at this level. Recall, this was explained theoretically in chapter two. If the same values are considered for level zero – from Table 64 - the horizontal binary question distance for the complete adjacency matrix is 2.0000 (there are four variables) questions; that is, at this level, because of the reduced number of variables created through the aggregation of data in the hierarchy, for a reduced number of level zero processes that recognises the reduced business scope in this scenario, it will take this number of questions to determine the state of the business. At this level, the micro analysis (Table 69) gives a value of 7.69899, which means that while it took fewer questions from the horizontal/macro perspective, this leads to a greater binary question distance from the vertical/micro perspective. All of which is consistent with the analysis of the basic model construction.

An important observation from the analysis of the basic matrix was that hidden information values for the subsets are similar to the value for the whole matrix; i.e. it would take a similar number of binary questions to identify the state of a subset, as it would the whole matrix. For this scenario the values, taken from Table 69 and Table 64 are provided in Table 70.

Table 70: Hidden information for level zero at macro and micro level.

	Overall Hidden Inf'	Plan	Source	Make	Deliver	Return	NGO
Level Zero Micro HI	7.68990	0	0	7.82018	7.98868	8.09276	6.85798
Level Zero Macro HI	2.0000	0	0	0	0	0	0

The earlier analysis of the basic construct resulted in an average 7.65240 additional binary questions to get to the state of the system after asking, on average, 2.58496 macro questions. In both scenarios, for each of the macro subsets, there are no additional questions; there is only one state for each subset at the macro level. Allowing for rounding, the micro value for the complete matrix should approximate to the average of the values for the subsets; in this scenario the results show, on average, 7.6899 additional binary questions to get to the state of the systems after asking, on average, 2.0000 macro questions. Allowing for rounding, the micro value for the complete matrix should approximate to the average values of the subsets. In this scenario the average value for the subsets equals 7.6899, matching the complete matrix value. The NGO values follows the same logic with the result that values remain the same irrespective of the level in the matrix, simply because there is only one NGO category at each level in the matrix; consequently the horizontal/macro value is always zero and the vertical/micro value is always 6.85798 in this case.

The key determinant of the entropy values is the reduction in the number of variables. In the basic analysis the process of aggregation reduced the number of variables from 186 at level three to 69 at level two and then 11 and 6 at levels one and zero respectively. In this sub-contractor scenario those values have reduced to 129, 41, 7 and 4 respectively. Entropy continues to demonstrate a micro level increase as aggregation increases, following ($S_d = S_o - S_a$). In this scenario the actual values are generally higher for the subsets due to the decrease in the number of variables and the additive nature of entropy: The k_{out} degree probability for each value and grouped sum increases due to the decrease in the number of variables, which reduces the probability values.

Graphically the values for hidden information and entropy are shown in Figure 55.

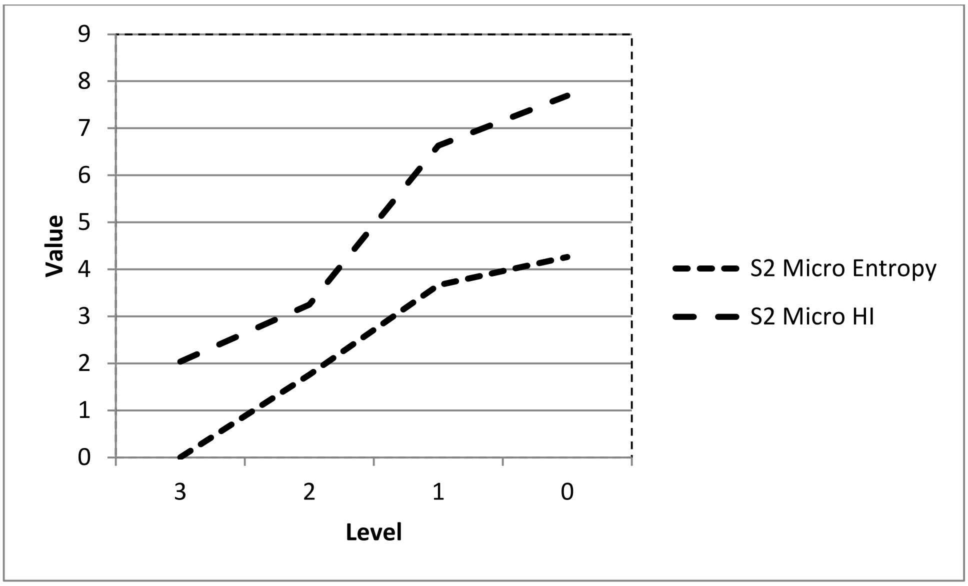


Figure 55: Entropy and hidden information 'micro level' for the subcontractor scenario.

The Pearson correlation co-efficient for the data= 0.98087; thus, assuming an alternative hypothesis to be there is no statistically significant correlation between entropy at the micro level and hidden information, and the null hypothesis to be that a statistically significant correlation exists. Pearson product moment correlation co-efficient: $r_{obt} = 0.98087$ compares to a two tailed test value for alpha = 0.5 $r_{cv} = \pm 0.9500$. Thus $r(2) = +0.98087, p > .05\ (two - tailed)$, demonstrating statistical significance between the two values.

Consistent with previous analysis, comparing the graphs for the horizontal/macro analysis (Figure 54) with those for the vertical/micro entropy analysis (Figure 55) shows the slopes for both hidden information and entropy to be inverse. For horizontal/macro analysis, as data is aggregated, values reduce; whereas, for vertical/micro analysis, as data is aggregated, values increase. Intuitively, this makes sense: From a vertical/micro perspective, the more aggregation, the greater the binary question distance. From a horizontal/macro perspective, the more aggregation, the lesser the number of binary questions. Conversely, a reduction in aggregation reduces the number of vertical/micro binary question and increases the number of horizontal/macro questions. A more detailed comparison of the slopes for each method and scenario will be the subject of the next section.

This section has analysed the adjacency matrix for the subccontractor scenario, at both the macro and micro levels. Consistent with previous scenarios, the macro analysis continues to demonstrate how, simply by constructing a hierarchical process structure and attaching K_{out} links, entropy and hidden information reduce from their consistent, equiprobable values. The analysis also demonstrates the correlation and difference in entropy and hidden information values for the scenario matrix. The micro analysis continues to demonstrate and validate the concept of question distance as average additional necessary questions to identify a given state structure; and how entropy at the non-aggregated process level will not 'see' this distance. The analysis continues to demonstrate how entropy simply reduces with the number of variables with figures that are unquantified. Finally, the graphical analysis shows the correlation between entropy and hidden information to be statistically significant and the slope of the graphs for micro analysis to be inverse to those for macro analysis; demonstrating again the two dimensions of entropy and hidden information that have been the subject of the research so far. The subcontractor scenario shows results that are consistent, with respect to entropy and hidden information, to those of the basic and prime contractor analysis. And it shows how the hidden information measure continues to provide definitive and measurable information. Through the addition of this scenario, the analysis, throughout, shows entropy to be only and indicator of 'mixed up ness', and not to be a measurable value of questions or question distance. Further, it shows how hidden information does not follow the additive restriction and does appear to provide values that are a true reflection of the number of questions required, both horizontally and vertically, to find an actual business state. The analyses thus far have covered basic, a multi organisation – prime contract, and subcontract scenarios. In these analyses we have seen how the graph slopes are inverse for macro and micro analysis, and how hidden information as a hypothesis compares with entropy. We have seen how the graphs for each slope in relation to the number of variables at each level of analysis; but, as yet, we have not analysed the rate at which the results change for each level of analysis. The next section will look in more detail at the slopes of each of the graphs.

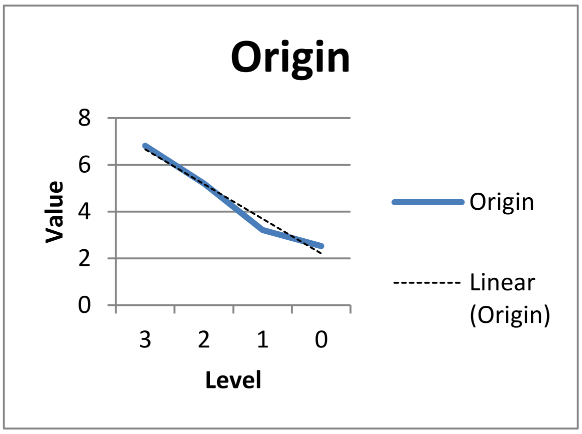
5.5 The Comparative analysis of each Scenario

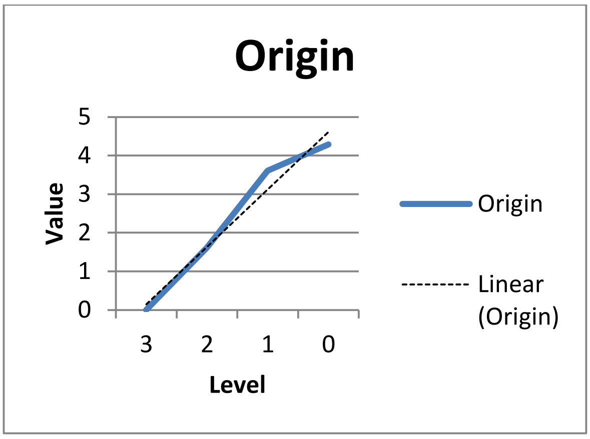
In this section we analyse the degree to which each of the graphs for macro and micro analysis slope. Firstly, six graphs are presented for entropy; followed by a further six for hidden information. Each graph plots the results from macro (on the left) and micro (on the right) analysis beginning with the original model; followed by prime contractor and sub-contractor scenarios. The values for each plot are entropy (followed by hidden information) for each level of aggregation - the same values in the above analyses. In these results a trend line and formula have been added for each of the scenario/macro – micro plots.

Below are the six graphs for entropy covering the macro analysis (on the left) and micro analysis (on the right). For each analysis there are three graphs covering the origin, prime and subcontractor scenarios.

Macro entropy analysis

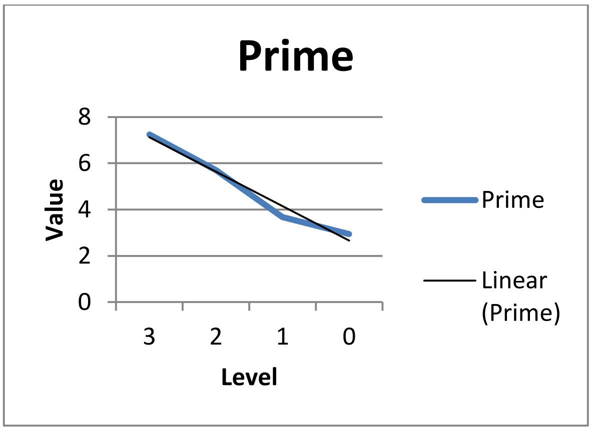
Micro entropy analysis

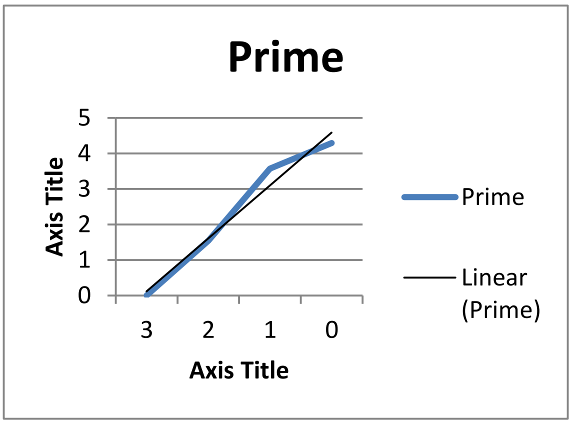




y = -1.4847x + 8.1492

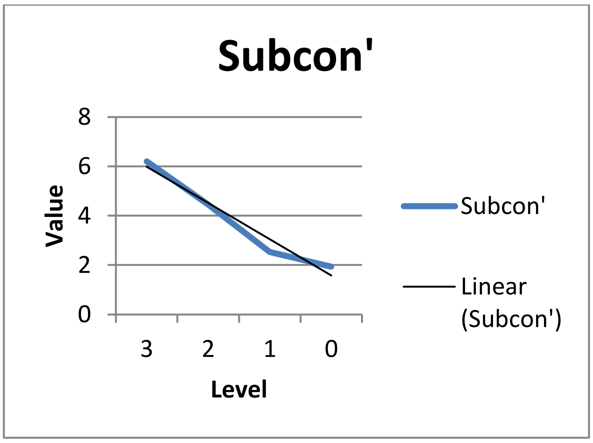
y = 1.4847x - 1.3357

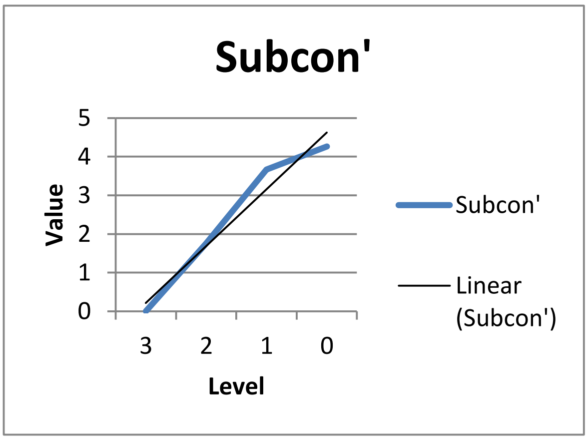




y = -1.492x + 8.6216

y = 1.492x - 1.3782





y = -1.4705x + 7.4563

y = 1.4705x - 1.2546

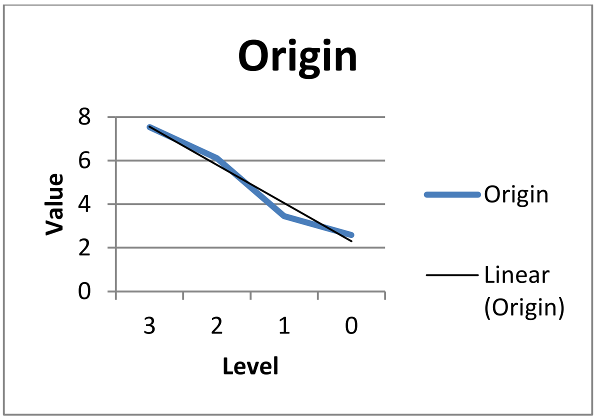
The two key points to note with the entropy graphs are: firstly the same but inverse values of a making the slopes of the horizontal/macro and vertical micro analyses to be the same for each scenario. Secondly, summing b for each of the basic, prime and subcontractor scenarios equals entropy for the overall matrix for the respective scenario, thus:

$$b_h + b_v = S$$

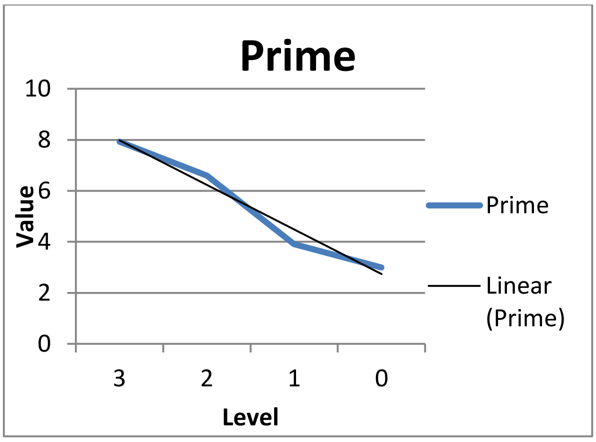
Equation 79

Entropy demonstrates symmetric characteristics, with absolute values of a equal, and the sum of b values for macro and micro b values equal to the overall entropy value for the scenario. Hidden information does not have the additive characteristics that may support this symmetry. Below are the six graphs for hidden information covering the macro analysis (left) and micro analysis (right). For each analysis there are three graphs covering the origin, prime and subcontractor scenarios.

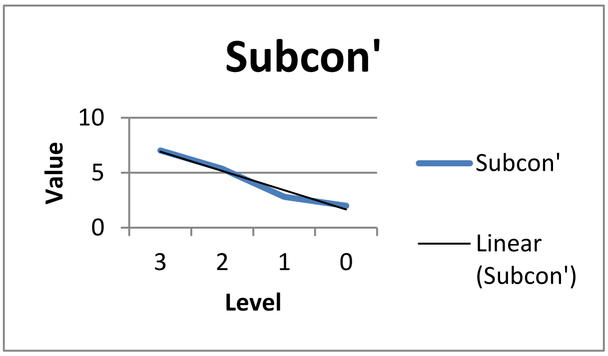
Macro Analysis



y = -1.7512x + 9.3009

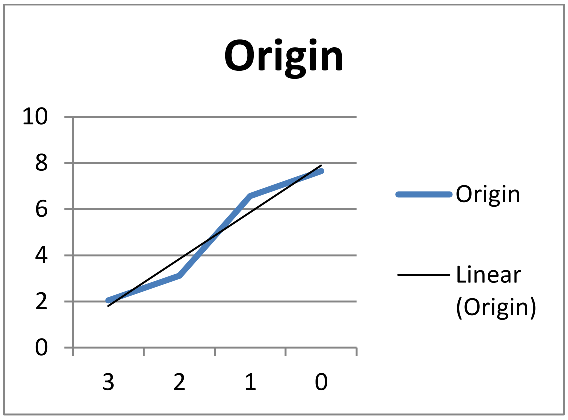


y = -1.7467x + 9.7248

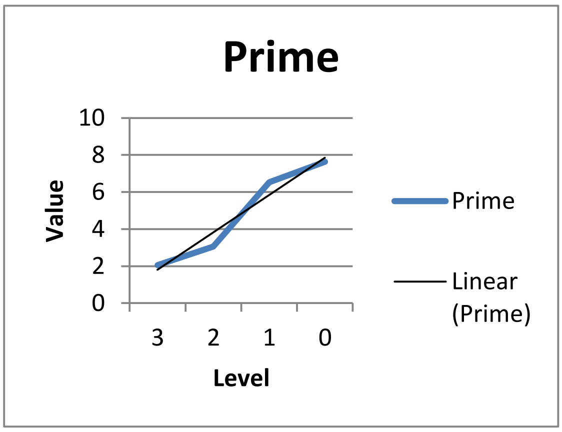


y = -1.7584x + 8.69

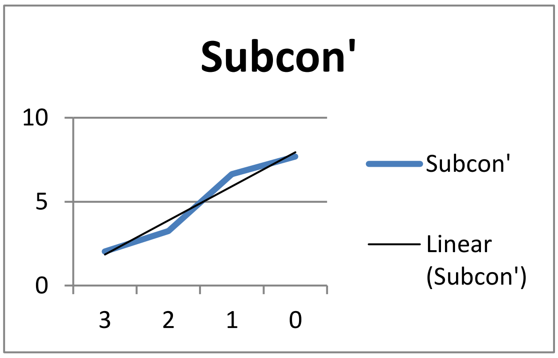
Micro Analysis



y = 2.0242x - 0.2134



y = 2.0181x - 0.2218



y = 2.0347x - 0.1855

Hidden information does not demonstrate the symmetry displayed in the analysis of entropy. But that is not to say there is no consistency in the results obtained. The following table summarises the a and b values for the hidden information graphs.

Table 71: Graph slope values for hidden information for each scenario.

	Macro		Micro	
	a	b	a	b
Origin	-1.7512	9.3009	2.0242	-0.2134
Prime	-1.7467	9.7248	2.0181	-0.2218
Subcont’	-1.7584	8.69	2.0347	-0.1855

For macro hidden information a represents the rate of change in $\log_2 n$ with respect to the change in number of variables over the three scenarios. Micro hidden information a represents the rate of change in the weighted sum of the $\log_2 n_m$ where n_m represents the number of variables in each group for each level. The numbers of variables in each group and level for each scenario shows how the structure changes with each scenario. The structure, basic or as one of the two scenarios impacts on the extent to which entropy and hidden information deviate. The next section explains this effect using and analogy.

5.6 The Structuring effects on Question Sets

Now that the results of the analysis have been explained, it is perhaps best to explain the relationship between structure and question sets by analogy and with reference to the triangular diagrams in Figure 9 and Figure 10. Imagine you are the director of a business function - supply chain management perhaps. You have six direct reports, each responsible for two sub-functions: Direct and Indirect Purchasing, Customer Management and Sales, Inventory Management and Logistics, Production and Operations Management, Budgeting and Finance, and Human Resources and Training. The managing director has asked you to get some specific information on the state of a major transaction, which involves you identifying exactly where the issue is in your organisation; using a binary questions process you would need $\log_2 6 = 2.58946$ questions plus one further question to identify the sub-function, 3.58496 in total. Now suppose each of the sub-functions has two further layers of governance; for ease we will assume these each consist of three capabilities, and each capability consists of five activities. As director, continuing with the binary question process, you would ask $\log_2 3 = 1.58496$, plus $\log_2 5 = 2.32193$ further questions to drill down through your organisation to get the the information you need; 7.49185 questions in total. Compare this with a scenario where your organisation is much flatter; where you have three direct reports each of whom has direct control over fifteen activities. In this scenario you would need to ask $\log_2 3 + \log_2 4 + \log_2 15 = 1.58496 + 2 + 3.90698 = 7.49185$ questions, the same as the first scenario. Now, let's compare this with the extreme scenario where you have no direct reports and access to all one hundred and eighty activities. Using the same process you would need to ask $\log_2 180 = 7.49185$ questions, the same again. Hidden information is consistent for all structural constructs using equal base quantities. If we compare this with the values calculated for entropy we see that in all three scenarios entropy would be the same 7.49185. This is because, in this example, the activities are all equiprobable. Equiprobability is where hidden information and entropy meet.

But in organisation structures, activities are not equiprobable; they are connected in the governance structure as a function of their in degree and out degree values. From a hidden information view the fact that activities are more or less likely is of no consequence, the same number of questions will be required to identify any given activity irrespective of its likelihood; the values for hidden information remain. But, for entropy the values change. For, instance, if one of the activities was 50% likely, and the remaining activities where equally likely over the remaining 50%, then the entropy values for the above scenarios would be 6.86816, 6.29869 and 4.76561 respectively which demonstrates that as structure deviates from equiprobability, entropy disconnects from binary question distance. The basic analysis explained the extent to which the supply chain operating reference model deviates from equiprobability by having a distribution of k_{in} ranging from 1 to 24, therefore, entropy in this model deviates from being representative of a number of questions.

And so, back to the example. We can see here, in simple form, one of the points being made in this research. You, as director and using the four layer example, ask an average of 2.58946 questions; after which you know only which of your direct reports owns the specific information you are looking for. The vertical micro component is the extent of the hidden information. Asking another question identifies which sub-function owns the specific issue, and the vertical component reduces further. The more you ask, the more you open up the granularity in the governance and the more detailed your information you have, which, in simple form, demonstrates the two dimensional measures - shown in Figure 9 and Figure 10 - required to understand how far, for a given governance structure you, as director, are from this detail.

In this example the organisation structure started with four layers, you as director and three other layers; the total number of activities was one-hundred and eighty. In the supply chain operating reference model there are the also four layer, with a few more activities; hence the analogy is not too removed from the actual values for layers and activities. Chapter five explained the extent to which the supply chain operating reference model deviates from equiprobability by having a distribution of k_{in}ranging from 1 to 24, therefore, hence the example – while being simple to understand - does deviate from being representative of the model under analysis; nevertheless the example is representative of the effect on the model under analysis and easier to understand.

This chapter has analysed the adjacency matrix construct. Beginning by setting out details of the baseline, subset and superset scenarios from the survey results analysis an adjacency matrix was constructed; the degree distribution for which was shown to correlate with classical degree distributions like, for instance, that of the world wide web. The degree distribution was also shown to approximate to the characteristic power law distribution, non-monotonicity for small k, and clustering a co-efficient suggesting a loosely coupled structure which was shown pictorially for the whole, and for each sub-element of the supply chain operating reference model. The analysis then set out the baseline position for equiprobability and explained the impact the basic non-equilibrium structure has on equilibrium values. The chapter then examined in some detail the baseline values for the macro horizontal structure at four levels of analysis and for each sub-element of the reference model. For four levels and six sub-elements (including the non-graph links) entropy and hidden information were analysed. The analysis showed how entropy, without units of measure and proving inconsistent with the number of binary question, begins to serve only as a measure of 'mixed up ness' for the matrix; whilst the two measure do correlate across all the sub-elements and the overall matrix. The analysis then moved to the analysis of the same four levels and six sub-elements of the reference model for the vertical/micro view of the matrix. This analysis shows how for a given level of macro question a vertical micro hidden component exists. The analysis also shows the additive and extensive characteristics of entropy begin to play against its use as a measure of questions; and how while different, the two measure correlate for the vertical analysis also.

To further validate the findings, the analysis continued by completing similar analyses for a prime contractor scenarios where planning and sourcing were carried out firstly as the prime contractor and then again for those elements of the business involved in manufacturing. This leads to analysis of four levels but, in this case, eight sub-elements. The analysis shows how entropy again does not represent question distances in the same way that hidden information does. Furthermore the analysis shows how entropy demonstrates the counter intuitive characteristic whereby entropy values reduce for each sub-element, suggesting some form of increasing certainty, when in fact the reduction is due to an increase in the number of variables. Despite the significant differences in entropy and hidden information the two measures continue to correlate throughout the prime contractor macro analysis.

The micro analysis for the same prime contractor scenario, using the same four by eight construct explained how entropy continues to demonstrate extensive, additive properties, while hidden information shows characteristics that average sub-element values and suggest how the structure itself creates hidden information distance as was pointed out in the basic analysis. The analysis also shows how hidden information measures respond to the increase in business scope with incremental increases in values, while entropy continues to demonstrate the counter intuitive reducing characteristic. Never the less the two values continue to correlate.

The analysis then moved on to look at the sub-contractor scenario where the planning and sourcing elements of business capability were removed and assumed to be completed by the prime contractor. The sub-contractor scenario consisted of four layers but, in this case, only four sub-elements. The horizontal/macro analysis and a comparison with the data from the basic analysis, further justified the finding that entropy only provides and indicator of mixed up ness and hidden information to provide a consistent quantifiable measure of binary question distance. Not only that, the analysis also justified the finding that the omission of probability from hidden information removes the inconsistency brought about by changing probability values as a function of the number of variables. The findings of common extensive and monotonic properties, and non-common additive property of entropy were confirmed in this scenario also, along with further validation of the correlation between the two measures.

The micro analysis for the same sub-contractor scenario, using the same four by four construct found that hidden information, being extensive and not related to probability continued to provide a quantifiable measure of question distance. The analysis also found that the issues associated with entropy continued to be valid in this construct; the extensive property of entropy when linked to probability, and with the absence of quantification leads entropy only to be an indicator of mixed up ness. And yet, the correlation between the two measures continued.

The analysis then moved on to look compare the slopes of the macro and micro values for the basic, prime and sub-contractor scenarios. For entropy it was found that the comparative alpha values for the trend line for macro and micro analysis were the same absolute values and that, for each of the scenarios, the sum of the beta values for the trend line equalled the entropy value for the overall matrix for the given scenario. Hidden information, on the other hand did not demonstrate the symmetry found with entropy, but it did highlight the effect the structure has on values, an effect that was analysed further.

The final part of the analysis, explained the impact the structure and stratification has on values. It demonstrated how entropy and hidden information deviate from the point of non-equilibrium and how the adjacency matrix used here was a non-equilibrium construct. Using an analogy, the analysis moved on to explain the operational effect – from a directors perspective - of the non-equilibrium, grouping and stratification effects on the distance directors may be from actual state values, and how, without understanding horizontal and vertical distances, this can lead to incomplete information. In the next chapter the findings will be discussed in the context of previous research in these domains.

6 Chapter Six: Discussion

This research was introduced suggesting supply chain management has in some industries struggled to gain recognition as an important aspect of business operations. Furthermore, current transitions in the economic climate, rising debt levels and spending reductions across the western hemisphere and growth in economies such as India, China and Brazil, all add to the globalisation and/or internationalisation pressures on organisations' supply chains. Internal to the organisation this pressure is further exacerbated; increasing levels of customer sophistication, mass customisation, the ever decreasing life cycle of most products, all add to the challenge faced by supply chain management. The combined internal, external, product and market challenges only add to the complexity of form in the supply chain necessary to rise to these various challenges. The introduction also suggested the consideration of supply chains as interacting structures or systems where, while valid, reductionist research approaches should be complimented with a new approach to considering this complexity of form. The suggested approach was to consider supply chains as systems of structure, defined by the governed process model of one or more businesses involved in the supply chain construct. With this form a measure of hidden information could be developed and compared with the more usual measure of complexity: entropy. Finally the introduction included a suggestion that, because of the only partial acceptance of the validity of a wholly reductionist approach, a thread from the grand theories of research methods should be included. Critical realism was the phrase associated with the grand theory selected.

The research hypothesised a quantifiable measure to be useful as a determinant of organisational complexity. The research hypothesis was built on the argument that there is an increasing importance on managing complexity in supply chain management and that increases in cost due to increases in complexity, without a concomitant increase in revenue, will be detrimental to the business; but managed increases in complexity that lead to increased revenue can be beneficial to the business. Hidden information was suggested as the measure of complexity in a supply chain and the hypothesis was that a quantification of hidden information provides a determinant of complexity of an organisations' supply chain that can be used to support supply chain governance and operational design.

This section aims to discuss the major findings of the research from both a content - recognising the laboratory experiment, and a process – recognising the grand theory thread – perspective. The major findings will be set out in the next section, followed by sections discussing meaning and importance, the relation to similar findings, alternative explanations, relevance, limitations and finally suggestions for further research.

6.1 Major Findings

This research has suggested a framework for the creation of a hidden information measure of complexity in supply chains can be constructed through an approach that brings together entropy, information theory and network theory to build an adjacency matrix and use an information theoretic measure to quantify the complexity of the governance structure in two dimensions. This section discusses the major findings of the research.

Earlier research, specific to supply chain management and its related subjects, has recognised the issues faced by organisations which include complexity, variously described. The proposal generally offered is that market pressures such as internationalisation, mass customisation and demand for increases in product variety are forcing businesses to adapt to more and more complexity. Generally this earlier research has argued that complexity can be measured using Shannon entropy, and the basis for measure has generally aimed to compare a specific aspect of the planned and actual state of the business. The conclusions from this earlier body of research can be summarised in terms of entropy providing a measure of this complexity.

The result from this research suggest entropy, without a unit of measure, without experimental result values that relate to binary questions, and when used in a single dimension, adds only an indicator value of mixed up ness for any given level of analysis in an organisations governance structure. The use of yes/no binary questioning, and measuring this using $\log_2 n$ - a measure more readily associated with Hartley's somewhat overlooked research - suggests a quantifiable measure of complexity as distance. The use of this measure overcomes the problem of units of measure, and provides results from this experiment that are consistently representative values of binary question distance. Furthermore, when combined with the creation of a framework that derives an adjacency matrix for a given governance structure, this measure provides a two dimensional perspective of complexity. The findings from this research bring into question assertions that entropy provides a measure or proxy for complexity in organisations. Other than mixed up ness, such assertions miss the issues highlighted above which are fundamental to informing organisations.

Literature more sharply focussed on entropy recognises the issues of units of measure and the validity of the Shannon entropy form; this body of research also questions the extensive property of Shannon – and other forms of - entropy. The unit of measure issue aside, the results herein support the proposal that the extensive characteristic of entropy is an issue: with the number of variables being analogous to volume, increases and decreases in volume produce a concomitant increase or decrease in entropy. At face value the extensive characteristic is valid. As volume decreases uncertainty decreases; but a question must be asked regarding the decrease in the number of variables that produce more certainty and less entropy. Is the reduction in volume – the number of variables – a cognitive, or observational, or exclusion decision? If so, then a reduction in entropy is aligned with a reductionist perspective; the process of reducing variables does not make these variables disappear; they are simply selected as variables to ignore. The process of exclusion to ignorance of a set of variables in the reductionist sense hides information and it is this information that is highlighted in a two dimensional approach used here. Highlighting this point appears to be unique to this research. Furthermore, and perhaps more importantly, this approach offers an alternative solution to the extensive property critique: The two dimensional approach counteracts the problem whereby entropy – and the hidden information measure proposed here - reduces with volume. As volumes reduce, macro level entropy and hidden information measures also reduce; but the two dimensional approach leads to a concomitant increase in micro levels values as the macro values reduce, the issue of extensive property is therefore removed.

The entropy and hidden information reducing – reductionist - process highlights the exclusion of information principle. This research shows the same phenomena is observed with the aggregation of process state information within an organisations governance structure. To go right back to chapter two, the observer may obtain the value seven on the role of two dice, but the observer cannot be sure of the value of each dice without further questions. From this research, the same can be said about observing aggregated information in a hierarchical organisation structure. In such structures, information aggregation 'hides' information as aggregation occurs through the structure. In such cases entropy does not provide meaningful information, except for mixed up ness as highlighted above.

Thus far differences between entropy and hidden information have been highlighted as units of measure and quantification through binary question distance. Both measures have highlighted a need for two dimensional analysis to overcome the extensivity critique. The results in this research show how a use of the hidden information measure provides a unit of measure, a more reliable quanitification of the unit of measure values and a set of values that are more consistent than the equivalent entropy values. While there are differences between the two measures, the correletion between them is high throughout the research; demonstrating how the use of the alternative method looses little when considering a measure for complexity in organisations, the analyst still gets a view of mixed up ness using hidden information. Thus the use of hidden information is more reliable and valid for measuring structural complexity, but it does not deviate wildly from the principle of entropy measured mixed up ness.

The literature on network theory explains how networks are constructed and measured in various ways. It also explains how network theory can be applied to supply chain management analysis and how entropy has been used previously in some forms of analysis. Standing out in the existing literature is the work of Easely and Klienberg (2010) in that they refer to the need for breadth first analysis of networks. A major output from this research is the specific application and validation of the breadth first concept to create a stratified analysis, and the development of a measurement concept for breadth first analysis that can be used in the analysis of supply chain management structures in this case, and any process structure in the general sense, from which an adjacency matrix can be constructed.

With the exception of Easely and Klienberg (2010), breadth first analysis remains fundamentally under researched. The findings here show – be it through entropy or hidden information – how ignorance of breadth first can mislead an observer. As macro entropy or hidden information is perceived to reduce, the micro component increases showing how information is being hidden from the observer because of aggregation of information on the granular state of the structure. Remember how the aggregation of blue and yellow sand produced – with the removal of granularity – a green state; the findings herein demonstate the same effect with structure aggregation. Another major finding therefore is the importance, in supply chain structure analysis, of understanding the extent to which information on the state of a structure is aggregated or consolidated within the governance structure before acting on that information i.e. the manager should a) understand the information provided and, b) understand the extent to which the information has been aggregated from raw state information.

In addition to the findings from the content of the research, there are findings from philosophical and critical realism perspectives. The next few paragraphs discuss these findings.

The method, set out in section five, proposed a laboratory experiment with a 'thread' of critical realism. The construction of the experiment; the use of network theory, the application of a breadth-depth analysis and creation of a measurement based on degrees to which information is available is in itself an example of specific information reduction to the point where the boundary conditions for the experiment exist. That is, in this research, specific elements of network theory and information theory have been selected from the set of all possible elements of network and information theory. The link to the use of hidden information is simply the creation of a triadic closure within the network of linked components that form the boundary condition of the research. Thus the research is an example of a generalisation the specific it is trying to research. Let’s explain this point again: The content of the research shows how a two dimensional approach measures the extent of hidden information as aggregation occurs. The process of completing the research selects applicable theories from a wider set of theories and completes network triadic closure on the chosen theories. Thus the reductionist process of completing the research is further evidence of the existence of the content.

The process of completing the research is also supportive of critical realism and an example of Bhaskar’s assertion that natural laws exist to be discovered and it is the closed experiment, designed by the epistemological drivers of the experimenter, which creates the empirical grounds for an argument of existence. This is what has happened here; the natural laws of entropy, logarithms and networks existed without this research. This research simply creates a set of boundary conditions – through the process defined in the last paragraph – to allow an ontology to be available for analysis. In this research chapter four took some time to explain the issues of research and develop a specific research paradigm. It is key to this – and probably any – research that this activity is undertaken to set the epistemological construct from which the research ontology is created.

6.2 The Meaning and Importance

In this research, the review of information theory and the works of Hartley and Shannon, explained the differences between the two; to a large extent this work has been a comparison of these two approaches: Shannon’s use of entropy and Hartley’s use of exponential – logarithmic values. Shannon’s works largely superseded Hartley’s and became the de facto measure in information theory. This research, along with other research, brings into question the wholesale adoption of Shannon’s entropy measure; suggesting that Hartley’s work, rather than being a precursor, should be considered as having independent validity that is worthy of further consideration and research.

Deriving an extension of Hartley's work, and applying this to a structured approach and stratified model of supply chain governance processes, highlights how entropy may not be as valid or reliable a measure as previously considered for it to be viewed as a measure of complexity. The findings here mean that hidden information as a quantified measure, set in the context of a stratified model, provides a more meaningful view for the supply chain manager on the complexity of the structure he or she has to manage. Furthermore, applied to the model, the approach defined here means the manager has a view on the extent to which he or she is distant from the information available for decision making. This is important because:

1. It provides clarity to the manager on the structure of the organisation.
2. It provides the manager with dimensions against which to view the information at hand.
3. It provides a measure of distance – which may be viewed in industry as degree of aggregation – which provides context for the information to hand.

This is important to any manager because it provides a measure of the balance of the information available. For instance, in a scenario where there is a choice between two pieces of information. The first piece of information is very precise and detailed; the second piece less so. An immediate response might be to simply view the information and be tempted to simply make a decision. Having a question distance to hand would suggest a request for more information on the second scenario before a decision is made. Having this balancing view available to the manager is an important output from this research.

This is not to say that entropy is unreliable or invalid; entropy does have a place. But the wholesale – or generalised - adoption of entropy is questionable, starting with the adoption of the term 'entropy' to describe a formula for probabilities. The terminology used in statistical thermodynamics clearly explains the need for the formula to be prefixed with Boltzmann's constant so that the term is distinct from the probabilistic formula $k\sum(p\log p)$, is different from $\sum p\log p$. Without the constant $p\log p$ is a formula, like sine, cosine or tangent, to be applied to a specific situation. It is important that this distinction is understood and clarified when applying entropy to a given business scenario or research hypothesis, and that an explanation of the equivalent of Boltzmann's constant be provided to apply the general formula to a specific research scenario.

One of the significant findings in the research is the suggestion that a two dimensional approach to observing entropy or hidden information is more informative and provides a construct for addressing the extensive property critique. The application of a two dimensional construct appears to be unique to this research. The extensive property critique is applicable to research that applies entropy or hidden information in one dimension; hence, when applying entropy or hidden information measures, research should address the extensive property critique challenge to add to the validity of the measures produced. Alternatively, research needs to enhance validity by explaining how only a single layer comparison is applied to any data.

The existing literature broadly relates complexity with variability; that is, variability is a component of the complex environment. It is. This research looks to separate out variability from complexity. The logic for this is that variability is an everyday occurrence in business, built in to the fabric of the governance structure; the management of which is a pre-requisite of business operation. Variability has an impact on the parameters of process execution and, as a consequence, the operational complexity. It does not have an impact on structural complexity unless it exceeds the structural components of the supply chain; in which case the structure and operations become chaotic; a higher order complexity. The approach and framework created through this research enables an understanding of the extent of supply chain process and parameter dispersion necessary to govern and operate a supply chain. Consumers are becoming increasing demanding, mass customisation is the operational norm, markets are global or international, or both; hence there is an emerging but clear need to be able to distinguish between variability and operational complexity, and structural complexity. The analysis of a supply chain governance structure through the creation of an adjacency matrix that captures the stratified characteristic of that structure, and the creation of a two dimensional, quantifiable measure means a framework can be built for meaningful analysis. Supply chains are under increasing demand to operate in these complex environments. Adding to the body of knowledge in this area could, by extending already popular models, be timely in terms of business interest and application.

In constructing this framework the research has pointed out the different contributions of Shannon (1948) and Hartley (1927). The results from this research strongly suggest Hartley type measures – applied two dimensionally to a framework – provide a more valid and reliable outcome when compared to Shannon's approach. With this and more general critiques of entropy, this research strengthen the case for a refresh of the generally assumed basis that Shannon's work should be the de-facto standard for the application of 'entropy'. It is important that – in future – research challenges the de-facto standard, differentiates between the analysis of future probability, and perhaps looks in more detail at the application of Hartley's principles.

Given a general adoption in this research of Hartley's principles, an important factor in the findings is that compared to entropy little is lost; the values for the two methods correlate closely. In doing so, the hidden information measure provides quantification, stratification and a framework for two dimensional analysis. This coupled with the close correlation suggests the hidden information approach to be much more useful.

The literature on network theory in supply chain management, of alternative approaches to network theory and the application of entropy and information theory in supply chain management, highlight the need for greater understanding of complexity in supply chain management if the concept is to remain an effective model for understanding the structure of a multi-organisation approach to the movement of material, information and cash. Furthermore, the literature highlights the need for new approaches to methods and frameworks for understanding complexity in the supply chains. Finally, the research calls for continued experimentation and testing of methods and frameworks in what appears to be a relatively under developed research field. Research on the use of breadth first analysis in networks is more limited and yet the approach lends itself to structural analysis and the identification of organisation state. The two dimensional approach adopted in this research highlights the importance of a breadth first concept of analysis so that the stratification in the structure can be identified. Stratification in the structure can be a significant issue for a business.

Chapter five commented on hidden information and business functional silos. The hidden information approach highlights the silo construction of business functions: the non-additive nature of hidden information shows how question distance for one function can be very similar to question distance for the whole business. The implication is that asking questions of a particular function is just as time consuming as asking questions for the whole business, which presents a powerful argument for a supply chain management approach to business governance. Integrated non-function governance constructs – like those of a supply chain – are easier to understand than the traditional silo’d business structure.

Philosophically this research is important because it brings into focus the issue that a reductionist principle in research methods suffers from the same extensive property. A reductionist approach – as is adopted here – manifests a reduction in entropy and hidden information as variables are cognitively or observationally reduced. This is important for three reasons:

1. The research is not immune to the characteristics it aims to address,
2. The process of completing the research is evidence in the research domain of the consolidation, aggregation, summarisation, process the research is addressing in the supply chain management domain,
3. The research highlights the need to address the extensive principle issue within a research methods approach.

Furthermore, it is important to recognise this philosophical point to be supportive of the critical realism thread.

In this research, the laboratory experiment, plus critical thinking, lends itself to a methodology based on network focal closure where nodes (actors in a system) develop triadic closure on a focus point. That is, the point, issue at hand, subject of analysis, is discussed and argued and entropy or hidden information as the available states of the subject of discussion gradually reduces. The network of possibilities is reduced or - using critical realism terminology- made transient, to the extent that focus alignment and triadic closure is achieved. This argument correlates with Bhaskar's introduction to his text on critical reasoning wherein the basis of assumption is that laws exist in nature - the Higgs boson was there before the experiment that proved it. The theory, derived by the experimenter, represents an occurrence of the natural law that can be associated with a closed set of conditions - known as the experiment

- which constitute the empirical grounds for the argument. Bhasker's argument is there is an ontological distinction between the two and, as such, the argument put forward for the construction of the critical realism perspective is far greater than the outlines suggested in some texts on business research. Critical realism is a far more fundamental case for a different philosophical perspective on business science that argues the constant conjunction, claimed necessary and sufficient for causal law, is neither sufficient nor necessary for a law to exist. For example, natural law requires the Higgs boson to exist. It exists, and would exist, without the intervention of man. Man contrived the existence of the boson and created a closed system experiment to prove the existence. Thus, the experiment is a specific closed ontology for the Higgs boson created by the experimenter, and the experiment is neither a necessary or sufficient condition for the existence of the boson. This is the case in this research: hidden information exists in business and this research is not necessary for the existence of the phenomena. The process of reducing hidden information (or entropy, if you prefer) from the content perspective, and the research approach perspective is not necessary for the existence of the phenomena; they are necessary for the experimenter's construction of an epistemology and ontology to demonstrate the existence of the natural law. This shows the importance of addressing how research brings together the process of undertaking research as reductionist extensively principled natural phenomena, with the subject as a separate natural phenomenon, and the principles of critical realism.

So there are three important points of triadic closure in this research. Firstly, there is the focal closure on the issue of the application of an approach and framework, and the use of the hidden information instrument being a more reliable and valid measure of complexity. Secondly, there is the triadic closure of the point that the process of conducting the research is not immune to the extensive property critique and as such is empirical evidence – from a separate domain – of the research phenomenon. Thirdly, the research topic and the research approach serve as evidence of the validity of critical realism as a grand theory that is – perhaps – less of a grand theory, and more of a more detailed method of creating understanding. Given this the next section will look at this research relative to findings from other research.

6.3 Relation to similar findings

In this section we will look at the findings from this research and how these findings related to findings in related research domains. Research using entropy in subjects not specific to, but in subjects related to, supply chain management (Yao, 1985; Kumar, 1987; Kaput & Kesavan, 1992; Sivadasan, et al., 2002); and the research specific to supply chain management (Blecker et al. 2005; Frizelle & Woodcock, 1995; Calinescu et al., 2000; Isik, 2010.) has distinguished, in some form or other, between structural and operational analysis. In these cases structural analysis has tended to lean towards linking structural complexity with the planned state of the business, and operational complexity with the actual state. In comparison, this research compliments previous work in this field by measuring the governance structure of the supply chain from which the planned and actual states of the business emerge. Throughout this body of research entropy is the prime measure. Ebrahimi et al. (1999) recognised some of the issues with entropy and chose to define entropy as a measure of diffuseness of buiness process. The findings herein support their assertion. Similarly, the findings here do not dispute the appropriate use of entropy; however, it does call for more clarity on what entropy does and how it can be applied. As a relevant aside, Lee Smolin, a professor of theoretical physics at the Perimeter Institute for theoretical physics, proposed methodologies for conceiving of a set of questions that identified the quantum state of a system, Smolin (2013). The methodology proposed here aligns with the logic used by Smolin in the quantum mechanics domain. Smolin argued for the volumetric relationship between question sets and the degrees of freedom in the quantum system. The method used in this thesis follows similar lines in that the number of binary questions required relate to the degrees of freedom in the organisation structure as a system. In his text on the issue, Smolin has no reason to refer to the abstraction of information, but his text does highlight a point that is evident here: Smolin explained the difference between the world of classic mechanics, a world of determinism and predictability; and the world of quantum mechanics, a world of probability. He theorises that the degree of freedom afforded to the quantum world can be understood by understanding the amount of information needed to predict some future state. Smolin’s assertions are, principally, the same as those described herein except for the addition here of the challenge of abstraction. If considered, the challenge of abstract would present itself in Smolin’s ideas; that is, the amount of information needed to predict a future quantum state would be less if information had been abstracted into an extensivity principled, entropy reducing hierarchy. At this point the concepts would align, Smolin’s relationship between questions and states in quantum mechanics would be the same as the question setting explained herein for supply chain architectures. Research specific to entropy (Ben-Naim, 2008; Tsallis, 2009; Li & Vitanyi, 2008) has raised doubts over the use of entropy as a de-facto standard. The findings in this research support these doubts by showing how entropy only provides limited information when other measures which correlate with entropy, provide more valid and reliable outcomes.

In chapter two, the application of entropy to supply chain management, Sivadasan, et al. (2006) differentiated between fine and coarse grain measures, arguing - because it allows for the seven, plus or minus two, states suggested to be the limits of human cognition, and discounting fine graining as not of interest for day to day high level management – for coarse grain analysis. The findings from this research offer a number of points with regard to these assertions. Firstly, the need for coarse grained information is one of the drivers for the abstraction process used to create meaningful information; that is, Sivadasan, et al. (2006) argue information is only seen as meaningful when it is abstracted to the extent that it is within the boundary of human cognition; information has to be aggregated to seven – plus or minus two – pieces of information to be within the limits of cognition, it is the human limit of cognition that drives aggregation. Secondly, the argument in this research is that it is not that fine graining is not of interest in day to day operations; fine graining is of interest. But the degree of granularity sits outside the boundary of human cognition. In such cases, which are necessary for effective operation, the extent to which there is a distance between the course grained information used to make decisions, and the fine grain information used in the abstraction of the course grained information, is informative for the manager as the understanding of the amount of abstraction that has taken place in the derivation of the course grain information; how far the manager is from the detailed information when he or she uses the cognitively acceptable information to hand.

Chapter two described the differences between standard deviation and variance as measure, and entropy; where entropy was described as a measure of diffuseness and standard deviation and variance as a measure of average distance from a given mean (Dionisio et al., 2005; Ebrahimi et al., 1999). A suite of business processes defined by a governance set, is created to manage the set of business scenarios, with the pertinent business capabilities, to move material, information and cash through the supply chain. Some scenarios are likely to be more popular than others; receiving a customer order should be more popular than processing a customer return! While some scenarios may be more popular than others, the idea that scenarios and processes are distributed around a mean is unlikely. To be clear, using an example, of course, receiving a customer order is – hopefully – a popular process and a process that may trigger processes for inventory allocation, production orders and purchase orders. The customer order will be for one of a number of customers; similarly, the production order will call a specific routing and the purchase order(s) one or more suppliers and items or services. Given that the inventory, production and purchase processes are all triggered by the receipt and processing of a customer order, it is more likely that the information flow across the scope of the governed business process structure, necessary to cover the scope of business activity, is diffused across the structure. This diffused state aligns with the findings in chapter 5; consequently this research supports the notion that measures of entropy and hidden information consider business state information to be diffused throughout the defined business governance structure, and the governance structure to be diffused structure laid out over the organisation structure.

The ability to analyse structure was developed using graph theory, from which the adjacency matrix was created. This approach was taken to fill – to some extent - the gap in research in this area (Anand & Bianconi, 2009), and the finding generally align with previous work in this domain. The different hidden information values found here for each of the sub-elements of the SCOR reference model, and at each level of analysis, demonstrate the heterogeneous structure of the reference model in the same way that heterogeneity was demonstrated by Sole and Valverde (2009). They went on to suggest entropy to be a measure of average heterogeneity in the structure. The findings here did not address this question, but being additive, entropy did not demonstrate an average value of the sub-element values. Hidden information, on the other hand, did demonstrate average sub-element values. The two dimensional approach used here adds another dimension to heterogeneity and further research is required in this area. Heterogeneity may – or may not – relate to complexity, it does not follow that necessarily heterogeneity equals complexity. Lassen and van der Aalst (2009), and Isiks (2010), pointed out that unnecessary complexity adds to costs. The framework developed here provides an approach for understanding complexity as distance and is – to an extent – theoretical and offering insight into a developing research domain. But, like most of the research in theis area, it is in need of testing and refinement.

The earlier chapters in this research, principally using the works of Easterby - Smith et al, (2012); Bryman and Bell, (2011); Hussey and Hussey, (1997); Barber, (1976); Stebbins, (2006); Malpas and Wake, (2013); Smith et al., (2009); Burrell and Morgan, (1979); Merton, (1967); Mingers and Rosenhead, (2004); Ferris, 2009 and Bhaskar, (2008) took some time to explain the approach to this research and the paradigm used and it is to this we now turn. In addition to the findings on governance structures, adjacency matrices, entropy and hidden information, the research is important when thinking about approaches to research. Herein, the process by which an approach to understanding complexity in supply chain management has been determined, has been reductionist: the problem of complexity in supply chain management has been reduced to a hypothesis regarding the interconnectedness of a set of processes. Hartley's (1927) principle of exclusion, explained in chapter two, demonstrates how, from an information theory perspective, the process of reducing is completed by defining what the information excludes. The information theoretic process of exclusion manifests itself in the reductionist approach to the generation of the hypothesis. Thus, in addition to the subject of this research, to which Hartley's work is applied, the process of completing the research is demonstrative of the way Hartley's principle applies itself. The process of arriving at the research question excludes – hides – information. The extent to which information is excluded or hidden is a function of research epistemology and ontology. Hartley's principle is relevant to both the subject and the process of this research. To understand this point in more detail some reflection on the work of Easterby-Smith et al, (2012) and specifically the assumptions underpinning Table 23, is required.

The assumptions for a source of truth, set out by Easterby-Smith, Thorpe, and Jackson (2012), for each ontological classification, explain the how a source of truth is defined. These assumptions are reproduced in Table 72:

Table 72: Ontologies and assumptions about truth. Reproduced from Easterby -Smith, Thorpe and Jackson (2012).

Ontology	Realism	Internal Realism	Relativism	Nominalism
Truth	Single Truth	Truth exists, but is obscure.	There are many truths	There is no truth
Facts	Facts exist and can be revealed	Facts are concrete, but cannot be accessed directly.	Facts depend on the viewpoints of the observer.	Facts are all human creations

The approach used in this research was to consider supply chain complexity – initially - to be a wide ranging subject with different perspectives and interpretation depending on the viewpoint of the observer; the Nominalist/Relativist approach. Gradually, the approach used reduced this position through a process of exclusion and aggregation, to one of a specific hypothesis that could be tested in a laboratory experiment. In other words, the process of research, in this case, is entropy reducing and, consequently, creates hidden information as the process progresses. The process of completing the research has contributed to the findings. Of course, there could be alternative explanations for the findings herein, and it is to these alternatives that we now turn.

6.4 Alternative Explanations

In some of the findings there are alternative explanations. In this section we will review the significant alternative explanations for our findings.

The results highlighted how sub-element entropy reduced in the prime contract scenario, and increased in the sub-contractor equivalent sub-contract scenario. The results asserts the reason is the relative increase and decrease in the number of variable, reducing or increasing the probability; demonstrating the counter intuitive position that in a multi organisation construct, the complexity is reduced, and increased as the variable reduces. An alternative explanation can be constructed from a structural perspective. This is that, for instance, in prime contract scenario, rather than additional processes being added at the same level in the process construction, as was the case here; the prime contract processes occur before those of the sub-contractor, and hence higher in the hierarchy. It could be that additional layers are added, rather than processes added to existing layers. Adding to the structure in this way would leave the number of variables untouched and the concomitant values unaltered. However, this only provides an alternative explanation for the prime contractor scenario. In the sub-contractor scenario processes are omitted at given hierarchical levels, and in these cases entropy and hidden information do reduce. The alternative explanation then is that, for scenarios that add to the processes, care must be taken in adding processes to the appropriate – or to a new - hierarchical layer. The outcome of this careful appropriation would be that entropy would remain the same for the horizontal hierarchical layer, but vertical layers would be added, reflecting an overall increase in diffusion. Because it includes a probability function, it is entropy rather than hidden information that is impacted by this problem.

Planning activities, in reality, are not tasks done once, they occur repeatedly through the operating cycle of the business, either through the calendar year, through the project cycle, or continually depending on the characteristics of the industry in which the supply chain operates. One view would be that production progression is dependent on a set of tasks to be completed in succession like, for instance, a production order routing. Correctly, to understand the state of a specific production order, a set of binary questions could be constructed against all the production tasks to determine the state of a production order; thus, in this case, the binary question set is associated with the state of a specific order over time. An alternative view is that the organisation is a value producing entity that is in all states –to a greater or lesser extent – always; each task occurs again and again to satisfy continual customer demand, creating continual value. For instance, a project engineering business – the building of a new ship perhaps, may want to identify the state of a specific project or sub-project; whereas an automotive manufacturer is in a set of distributed states, processing multiple vehicle orders, as a function of takt time. The distributed states view has greater alignment with this research and would allow for prime contract activities to occur parallel to the other activities in the business. The validity of these perspectives is dependent on an individual's position and role in the business, their epistemology and ontology of the business; their critically realistic way of conceiving the business. The project manager might be interested in the state of a specific production order, the programme manager might be interested in the resources demanded by a particular production activity for a group of customers, and the operations manager might be interested in the demand for resources and production effectivity – adherence to time, cost and quality, for the whole of the business. Generally, this research aligns with an operations manager perspective, specific to supply chain management capability. This does not challenge the project/programme manager perspective and additional research, applying the same principles to these ontologies may be justified.

So there are alternatives in need of further consideration or research. Given the general immaturity of research in this field and the various calls for addition research, it is not surprising that this research adds to knowledge and provides insight, and raises additional consideration. Nevertheless, this research has relevance and it is this that will be covered in the next section.

6.5 Relevance of the Findings

‘My greatest concern was what to call it. I thought of calling it information, but the word was overly used, so I decided to call it uncertainty. When I discussed this with John von Neumann, he had a better idea. Von Neumann told me: “you should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name. In the second place, and more important, no one knows what entropy really is, so in the debate you will always have the advantage’. Shannon’s discussion with Tribus in 1961 in which he quotes this conversation with John von Neumann explains how the label ‘entropy’ was associated with this ‘uncertainty function’. Be the story anecdotal, apocryphal or factual, the validity of the logic has been questioned by some researchers. This research is relevant because it continues to add to the body of knowledge that questions the validity of logic for using the term entropy. The research is relevant for a number of other reasons also.

The inclusion of the ‘thread’ of critical realism in three parts of this research (the content, the recognition in the methodology and the removal of the hypocritical critique by including a thread of critical realism in the text) is relevant because it has added to the richness of the output. This is a key point: time taken discussing the issues with research, and developing a research paradigm; while seemingly convoluted, have enabled richer outcomes, with the creation of two sets of findings beyond the findings of the laboratory approach. All of which goes some way to placate the criticisms towards the narrowness of academic research set out earlier.

The findings are relevant to business because they demonstrate how the more a business increases its range of processes and the scale of process variables, the more distant one may become from understanding the organisations state. As organisations struggle to cope with increases in complexity, management organisation structures and an aligned process governance structure, and an understanding of the distance the manager is from the actual state information is, through the use of a hidden information measure, a key dimension for understanding and acting on the information to hand. Management structures that assume a hierarchical model, if combined with a governance structure that assumes a more horizontal model; or the opposite positions where management structures are flat and the business structures are hierarchical, may see operational difficulty. Hence, the research is further relevant to business because it provides for an approach to the construction of a framework, using a business's governance structure – enabled by the business's process model and parameters – that can be utilised to measure the degree to which hidden information is created in the governance structure analysed. Business's work in quantifiable values: cash in a prescribed currency, profit in a prescribed currency, costs in a prescribed currency; each, weeks, days, headcount, customers, days sales outstanding, suppliers, inventory turns, creditor values, debtor values, etc. etc. The use of a two dimensional, quantifiable measure of complexity of a business's actual governance structure, and providing an approach and framework – that can be revisited or iteratively developed - aligns with business' need for quantification. On this basis the findings in this research are also relevant because they add more structure and knowledge in a business scenario. Finally for a business this research is relevant because it structures and quantifies an intuitive management perspective; 'We need to drill down into…'. The research explains how a business’s own governance structure can be used to quantify these intuitions. In addition to the business domain, the research is relevant to the academic domain.

From an academic view the findings are relevant to research firstly because they incorporate the horizontal and vertical concept into the academic literature and propose an approach that overcomes some of the earlier limitations associated with the topic. The combination of the horizontal and vertical analysis provides a concept that overcomes the extensive properties critique; that is, considering only macro/horizontal measures leaves open the critique of extensivity. Measuring the vertical component in combination with the existing horizontal component overcomes this issue.

Hartley (1927) has featured throughout this research. It was explained how Hartley’s work primarily acted as a precursor and informer of Shannon’s (1948) research which became the de-facto standard and baseline for large volumes of research activity. This research has shown the principles applied by Hartley, and developed in this research as ‘hidden information’, add to academic knowledge and provide significant insight on the measurement of complexity in this context. The ‘hidden information’ principles overcome some of the issues arising from the use – in previous research - of the Shannon approach. Hence this research is relevant because it challenges the de-facto standard and provides insight into another stream of research that revisits and develops Hartley’s approach.

It was not necessary for Hartley (1927) or Shannon (1948) to consider their findings against an organisations’ governance framework; whereas, in other research, the structure of the governance framework has been considered. Generally, in that research a uni-dimensional approach was considered, with the entropy – Shannon or otherwise – being variously described.

In this context the relevance of this research is threefold: Firstly, this research has explained and demonstrated the importance of a two dimensional perspective when thinking about organisation structures in the context of entropy or hidden information measures. Secondly, this research is relevant when considering the application of entropy in certain business situations; it acts as an introduction to the possibility that Shannon entropy is not universally valuable for understanding complicated business situations. This assertion is not as radical as it may sound: The literature review, and specifically, Ben-Naim (2008) and Tsallis (2009), also raise issues of the applicability and validity of Shannon entropy. This research merely adds a new dimension to those questions in the context of supply chain management. Thirdly, this research offers insight into an alternative approach and construct using a relatively simple, quantifiable, valid and reliable measure, derived and developed from previous work, which offers opportunities for further research in the field.

The principle whereby certainty increases as entropy is perceived to reduce, needs further explanation to determine its relevance. The principle is described as an increase in certainty... entropy is reducing, certainty must be increasing. There are two reasons for this and it is unusual for these two reasons to be differentiated in research. The first reason for observing this principle is through aggregation of state information across the structural stratification, as we have seen in the above analysis on the macro condition. The second reason is through more contrast within the same number of variables. Aggregation creates a perception of increased certainty by reducing the number of variables over which information is presented, which is the extensivity principle created by the governance structure; the hiding of information. Increasing contrast across the same number of variables does increase certainty in this respect. The two reasons are distinct, offer different reasoning, and should in research, be differentiated. The use herein of a matrix structure, the two dimensional analysis and the use of a more simple measure (hidden information excludes probability) differentiates between aggregation layers and compensates for the differentiation by measuring the vertical component. The research is relevant in this respect because it provides a more robust approach to complexity measurement in this context.

The distinction between aggregation and contrast is important and – for different and additional reasons - worthy of more reflection. Chapter four took some time to identify a research methodology; the discussion included a review of the work of Bhaskar (2008) who, under the banner of a justification for critical realism, principally distinguished causal law and empiricism. To support his arguments he specifically describes and differentiates between three determinisms: regularity determinism, which follows a classic if x, then y and which Bhasker criticises due to its ‘ontological presupposition’ and ‘restriction in methodological responses’; ubiquity determinism which assumes ‘every event to have a real cause’; and intelligibility determinism which assumes that every event has an ‘intelligible’ cause. He argues for a type of hierarchy covering the three determinisms. Anything that causes an event is ubiquitous; only those events that can be perceived by humans are intelligible, and those that can be assigned causal laws for a given boundary condition can be assigned regularity determinisms. Summarily this transition from ubiquity to regularity is a transition from equiprobable (known or unknown) events through to a regularity (certainty) where events can be ‘caused’ through the creation of bounded conditions by the researcher, thus ensuring research reliability and validity. The move from ubiquity to regularity is a transition through variable identification, variable reduction and aggregation, the assignment of variable values and the analysis of variable; which is an entropy reducing transition as described herein. Thus it seems a characteristic of the process of aggregation or variable reduction – the process of creating hidden information - may miss a differentiation between classes of determinism. The process of variable reduction – the entropy reducing, hidden information creating approach explained herein – can be considered as a model that supports research method definitions and particularly critical realism. This relevance to research philosophy will be discussed in more detail in the next few paragraphs.

Chapter one outlined the systemic approach to viewing organisations, primarily using four components of systems thinking described by Gharajedaghi (2011), and chapter four considered the use of Morgan's (2006) assertion that organisations can be deconstructed into nine ontologies, but these ontologies should be considered as views rather than definitions; all of which are concurrent in daily operation. The findings of this research are relevant to the position set out by both authors: Firstly, in support of Morgan's research, governance structures – supply chain management in this case – demonstrate how organisations as machines deconstruct into functional components and sub-components, and how these sub-components deconstruct into capabilities and processes. The business operates by transitioning and transforming products, information and cash through the processes activities and across the necessary links in the same or adjacent processes. The effort added by resources to make the process happen adds to the organisation as a system such that the organism metaphor appears relevant. Hence the analysis in this research of the structure is supportive of the metaphors proposed by Morgan. Secondly, Morgan explains all metaphors to be concurrent, an organisation behaves like a machine, and an organism, and a brain, and a culture, and a political system, and a psychic prison, and in flux and transformation concurrently. Gharajedaghi (2011), described thinking about a system using four components: holistic thinking, sociocultural modelling, operational thinking and design thinking (see Figure 1). The results in this research show how entropy, being additive, naturally treats sub-components of larger systems as being independent. Hidden information, on the other hand, is exponential rather than additive and the hidden information approach shows how it is more effective for operations management to think about the organisation as a system of inter-functional interactions – as a supply chain – rather than a set of interacting functional components. Think about it this way: sticking with the examples used throughout this research, if a given function has thirty two possible states it will take five binary questions to identify one of the possible states. If there are two functions with the same number of states, and they are managed as interacting functions, it will take $5 + 5 = 10$ binary questions. Whereas if the two functions are managed as a system of sixty four inter – functional interactions, it would take only $\log_2 64 = 6$ binary questions. Hence this research is relevant to the systems view because hidden information shows how the interacting systems can be more efficiently understood.

Another philosophical view is the findings explain the general tendency of individuals, organisations or other entities to coalesce around a given position. The decision or concept making process we all follow as we take on board as much information as we can, or want to, to make our own internal constructs of the world demonstrate this process of aggregating information, hiding information and reducing entropy in our cognitive world. The same can be applied to organisations who, as they follow the above process for all the information they consider topical, hide information and reduce entropy in their organisation. In chapter four, Morgan’s (2006) eight organisation metaphors where explained. The construct explained here suggests an additional metaphor; that of organisations as a systems of entropy reduction; taking material and information at a given level of reduced entropy, processing this material and information to a further reduced level of entropy and exchanging this for revenue (a car is the output of an entropy reducing process of converting raw material and raw information into a specific car; a consultancy point of view is information aggregated and reduced from all possible information available). The organisation as a mechanism of bonding, or organisation as certainty creator.

The thesis took some considerable time explaining the approach to research, the strategy, approach and paradigm, before arriving at an internal realist ontology. The research paradigm was explained in chapter four which described the experimental nature of the research and the laboratory approach taken to protect construct and predictive validity, and the reliability of the research if field data variables are included at such and early stage. There is a dimension to the research that reflects on the relevance of the findings in comparison to research methodologies; this section will look at this findings in this way. To do so the section will firstly look at the relevance to ontology, followed by relevance to epistemology. Both will begin with the assumption of internal realism discussed during the selection of a research paradigm.

The internal realist approach, (Table 72), explained truth to exist and be obscure, and facts to be concrete and not directly accessible. The findings in the research are relevant to the research perspective for a number of reasons: chapter four explained how adopting an either/or selection of positivism or phenomenology had been criticised, and that a tendency towards one or the other was preferable. The section also explained the argument proposed by Burrell and Morgan (1979): each research approach gives rise to different research outcomes due to the underlying assumptions. In this research we have seen how horizontal and vertical hidden information measures can be applied to a governance structure that aggregates organisation states. If we consider a scenario – similar to the prime and sub-contractor used here - where the structure is set out differently, different values for hidden information would be obtained. Now, assume for a moment that two structural scenarios are constructed from the same underlying base, the number of level three variables stays constant at 186, but the aggregation structure for the two scenarios is different. The horizontal and vertical hidden information values for the two scenarios would be different and different interpretations would be created. It is therefore very likely that the application of the approach used herein would be relevant to the validity of the assertions made by Burrell and Morgan (1979).

A multi-dimensional approach, or an approach using one of the so called 'grand theories' goes some way to answer the criticism from business that some academic research is 'too narrow' in its approach as a consequence of academic rigour. The inclusion of a defined set of researcher baseline assumptions concerning the purpose of research and its epistemology led to the pluralistic approach that in turn led to the term grand theories, of which critical realism is an example. We have seen in this research how aggregations in structures create horizontal certainty by hiding information vertically. In the scenarios above we suggested how different structures applied to the same basic values would create different aggregated information. The same view can be taken with regard to researchers' ontological constructs: the researcher sets out the initial boundary conditions, applies a process of aggregation or exclusion to arrive at a specific closed system necessary for research validity and from which a natural law can be determined. The creation of further scenarios would lead to further research that would further test the validity of any natural law; a process that aligns with that set out by Bhaskar (2008). If we allow for the approach, framework and measures set out in this research to be used as a general approach, then the research is relevant to the creation, application and justification of research ontologies; it asks the researcher to deliberately describe the initial boundary conditions, the aggregation, the exclusions and the final conditions from which the results are determined.

In the phenomenological construct, where the focus is more on the how a reality is constructed by individuals and language, the research remains relevant in that it lays out a process for starting with a wider ‘environment’ construct and through the process of aggregation, exclusion etc. a reality is constructed. In the context of ontology then, positivism is the outcome of the process and phenomenology the start point, process and end point. So it follows that the research is relevant to research methodologies because it contributes to an explanation of the way in which an ontology can be created.

Epistemologically the findings continue to be relevant because it is entirely reasonable to suggest that just as the approach and framework can be applied to the topic of analysis, so to can the approach and framework be applied to the way of knowing about the topic. After all, the discussion set out in chapter four moved from a wide boundary condition on research methods, through a process of aggregation or exclusion, to the application of the applied method. Furthermore, with the framework and approach described here being applied – knowingly or otherwise – to ontology and epistemology constructs explains how the double hermeneutic interpretation issue can be brought into the open through the creation of a framework and approach.

In this research, using an internal realist perspective, the process of research appears to be as relevant as the findings of the research. The process of research being a general example of the findings determined from a supply chain model; where a process of determining ontological and epistemological constructs create hidden information (primarily vertical) through a process of aggregation and exclusion.

In summary, the basis for internal realist ontology assumes the truth exists, but is obscure; and facts to be concrete but not accessible directly. In this research we have discussed how hidden information changes with the structure of the organisation. Hiding information differently and necessarily leads to different visible information and interpretations of that information. The research is therefore relevant to the selection of a research ontology and epistemology because, as a general approach, it supports the determination of how truth and facts are accessed.

The research is relevant to questions on future organisational constructs and how these may be understood. Recently, radical views on organisation constructs have been suggested, Hamel (2012) for instance. These suggestions include the idea that an organisation structure be turned on its head: What if connection between all possible activities in the business were ubiquitous; that is, all possible k's are available to be brought together on a case by case basis to address the needs of the business today, and tomorrow and whenever – this can be viewed as the Gilbert construct for networks, described in chapter two. Instead of aggregating and pushing information up through what is a restrictive hierarchy, the organisation construct has a greater number of processes, offering a greater range of services, that come together in a form best suited to customer demand…on this occasion; a sort of mass service-ation of the business. Understanding such constructs would be complex, and the creation and use of an appropriate measure would be key to ensuring the effectivity and efficiency of such a business. The approach, framework and measures used here offer some limited insight into how constructs such as those suggested could be measured; however, much more research in this domain is required and while this research does contribute, the example points out a limitation of the research in that it only analyses one governance framework. There are several other limitations to the research, these will be discussed in the next section.

6.6 Limitation of the Research

There is a close correlation between hidden information and entropy when using the supply chain operating reference model as the basis for the construction of the adjacency matrix and while it is likely that adjacency matrices would show similar properties, this cannot be claimed in this research. For instance, the Voluntary Inter-industry Commerce Solution model, could be used to further analyse the approach, framework and measures proposed herein.

The methodology used in the research was a laboratory experiment; not unusual for this field of research. And similar to points made in other research in this field, there is need to further test and validate the general concepts, and to move the concepts towards different methodologies aimed at testing in a business sense.

The motivations for this research, set out in chapter one, explained the systemic perspective taken by the author; and the development of a revised model, set out in chapter two, explained the intent to offer a descriptive approach. In addition, the research paradigm, described in chapter four, set out the experimental intent of the research and the laboratory, rather than field, research approach. While valid, the motivation, descriptive and experimental – particularly laboratory – do limit the research by excluding mathematical rigour in favour of explanatory text, and creating a very specific environment in which to observe the experiment; for instance, mathematically, a section on the proof for entropy and the mathematical relationship between this and the formula for vertical hidden information would have provided further mathematical rigour on the relationship between the two. Constraining - or aggregating, or excluding, or hiding, to make the point again! – the research to a laboratory environment, while appropriate for the research, opens up the possibilities of further research in the field and the application of alternative research methods to the research domain to test the validity and reliability of the approach in different contexts. The next section will set out some of these research domain topics.

6.7 Further Research

The research has highlighted the possibility that the entropy reducing – I use the term entropy in this sense as a descriptor rather than a reference to the mathematical formulation – or hidden information, systems perspective, should be reconsidered by conceiving of the complexity in an organisation to be two dimensional, where increases in one dimension are balanced by decreases in the other. Further research is required on this subject to:

1. Validate the measures using different governance structures.
2. Add to the mathematical rigour in the solution offered in this research; for instance, a mathematically rigorous analysis of the relationship between macro and micro entropy and hidden information, or a similar analysis of the relationship between the vertical component of hidden information and the formulae set out in appendix 0 for the proof of entropy.
3. Develop insight in this domain by testing the approach using business data.
4. Develop further insight in the domain by applying different research methods; for instance, an action research project using the framework and measures during a business transformation project.
5. From a research methods perspective, add to the insight on grand theories, or to the processes of defining ontologies and epistemologies, by discussing the application of the approach to the stratifying or excluding process necessary for research method decisions.

The reader will recall the explanation in the literature review distinguishing between static measures, usually represented as some form of planned state and dynamic measures, possibly measured as actual states. In such cases the dynamic state measures may be derived from actuals represented as a weighted graph. This application of this type of construct to the use of hidden information measures, and a comparison between state and actuals would take the research a step further into the operational domain and thus represents an opportunity for further research in the field.

This research has considered only one, relatively small, aspect of – perhaps – a thermodynamic approach to viewing organisations. Systems modelling and contingency theory describe aspects of a complex set of interactions or views, and their application continues to be discussed and justified. These approaches, by analogy, correlate with aspects of thermodynamics which, obviously, can be thought of in a much wider sense. The theories and characteristics of the laws of thermodynamics, while seemingly appearing to align with some of the issues of business, have only had limited exposure and testing in the business domain. Consequently, there seems to be a significant opportunity to further research the application of the laws of thermodynamics to the business domain. Incidentally, the same argument can be applied to quantum mechanics, quantum biology and physics in general.

The form $Qn = Log_2 N$- used herein for the horizontal measure - is the same as the form $D = \log_e N$ where N is the measuring dimension, e the scaling factor and D the fractal dimension or Hausdorff dimension in the study of fractals. Furthermore, in principle, there are similarities between the vertical structure seen here and the Cantor set principle, the graph slopes captured in chapter five and fractal dimensions, and the extent of granularity used in the adjacency matrix and the box counting method of fractal dimensions. The principles used to measure fractal or Hausdorff dimensions therefore seem to lend themselves to the study of a stratified organisational construct in organisations, and a comparison or further analysis of organisation structure using these principles may prove to be additionally informing.

7 Chapter Seven: Conclusions

This research has taken a journey through research methodologies, thermodynamics, information theory and network theory to hypothesize that a quantification of hidden information provides a determinant of complexity of an organisations' supply chain that can be used to support supply chain governance and operational design. Internal realism, with a thread of critical realism, provided the roadmap, and at the end of the journey there are a number of conclusion we can draw about the places we have visited and the road taken.

There is a need to understand complexity, and the many truths proposition from internal realism suggests how the task of understanding complexity is complex. This research has used a development of an information theory view to measure complexity; which has been compared with entropy as a measure; a measure that, to an extent has become a de-facto standard. This is just one method, there is a need for further research on the ontologies and epistemologies of complexity. In this respect, this research has only scratched the surface.

The creation of an approach based on the formation of a tabular construct, determined from a governance structure, provides insight into the governance construct; the network decay function, for instance. The tabular construct then provided a set of data that can be analysed in two dimensions. Using two dimensions demonstrates how, without cognisance of the orthogonal dimension, a manager could assume information to be granular to the extent that the information provided can be acted upon. It also highlights how a manager could make an inappropriate comparative choice simply by selecting between two sets of information provided from different levels of abstraction or aggregation.

The need to understand in the organisation structure the impact of these two dimensions is important to the operations manager. As usual in such cases, a quantifiable measure will be more tangible, informative and generally useable than some arbitrary, or at least dimensionless, value. For this research the proposed measure, binary questions, was determined as an inversion of Hartley's approach for determining an amount of information. Applied to this methodology, hidden information proved to be a valid and reliable two dimensional measure of complexity.

Hidden information was compared with Shannon entropy; the results from both measures demonstrated a high correlation throughout. As a measure of complexity entropy offers some insight into indicator values of mixed up ness of a given set of business processes; however, it remains challenged by fundamental issues that frustrate its use for more quantifiable measures and it seems reasonable to question the use of Shannon entropy as the de-facto standard. Hidden information and the two dimensional approach offers a methodology for overcoming some of the issues; hidden information has a unit of measure and the two dimensions of the approach offers an approach to overcome the extensivity issue found with a single dimension approach.

The purpose of this research was to add to knowledge by offering a revised approach, hence the structure is aimed at doing just this. In doing so, a number of further research avenues are identified: Firstly research aimed towards a more robust field experiment; and secondly, a more thorough understanding of the mathematical implications of the proposed approach. That said, the findings of the research do offer further insight into the subject; with a revised approach to quantifying complexity and by overcoming some of the previously identified issues. The hope is that this will motivate further research in this important domain.

Throughout the research a thread of critical realism was maintained. The maintenance of this thread has shown how the process of creating both ontology and epistemology of research follows a process similar to the aggregation and exclusion process identified in the two dimensional approach to creating hidden information. Consequently, the two dimensional approach to structures offers a view on how ontology and epistemology for any given state are formed.

References

Ackoff, R. L., & Emery, R. E. (1972). *On purposeful systems.* London: Tavistock Publications.

Anand, K., & Bianconi, G. (2009). Entropy measures for networks: towards an information theory of complex topologies. *Physical Review, 80,* pp. 1-4.

Ashby, W. R. (1956). *An introduction to cybernetics.* London: Chapman Hall.

Axelrod, R., & Cohen, M. D. (2000). *Harnessing complexity: organisational implications of a scientific frontier.* New York: The Free Press.

Barber, T. X. (1976). *Pitfalls in human research.* Oxford: Pergamon.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17,* pp. 99-120.

Barret, A., Barthelemy, M., & Vespignani, A. (2011). *Dynamical processes on complex networks.* Cambridge: Cambridge University Press.

Beer, S. (1994). *The heart of the enterprise.* New York: Wiley.

Ben - Naim, A. (2008). *A Farewell to Entropy: Statistical Thermodynamics Based on Information [Kindle edition]*. Retrieved from www.amazon.co.uk.

Bennet, A., & Bennet, D. (2004). *Organisational survival in the new world: the intelligent complex adaptive system.* Oxford: Elsevier, Butterworth - Heinemann.

Bhaskar, R. (2008). *A realist theory of science.* Abingdon: Taylor & Francis.

Blecker, T., Kersten, W., & Meyer, C. (2005). Development of an approach for analysing supply chain complexity. *International Mass Customisation Meeting 2005,* (pp. 47-59). Klagenfurt.

Bowersox, D., Closs, D., & Cooper, M. B. (2007). *Supply chain logistics & management* (2nd ed.). New York: McGraw - Hill.

Bryman, A., & Bell, E. (2011). *Business research methods* (3rd ed.). New York: Oxford University Press.

Buchanan, D., & Bryman, A. (2007). Contextualising methods choice in organisational research. *Organisational Research Methods, 10*(3), pp. 483-501.

Burrell, G., & Morgan, G. (1979). *Sociological paradigms and organisational analysis.* Aldershot : Gower.

Calinescu, A., Efstathiou, J., Sivadasan, S., Schirn, J., & Huatuco, H. L. (2000). Complexity in manufacturing: an information theoretic approach. In P. McCarthy, & T. Rakotobe-Joe (Eds.), *Proceedings of the complexity and complex systems in industry* (pp. 30-44). Warwick.

Cardoso, J. (2005). How to measure the control flow complexity of web processes and workflow. In *Workflow Handbook* (pp. 199-212). Lighthouse Point.

Choi, T. Y., Dooley, K. J., & Rungtusanatham, M. (2001). Supply networks and complex adaptive systems: control vs emergence. *Journal of Operations Management, 19*, pp.351-366.

Coase, R. H. (1999). The nature of the firm. In O. E. Williamson, & S. E. Masten (Eds.), *The Economics of Transaction Costs* (pp. 3-22). Cheltenham: Edward Elgar Publishing.

Cousins, P., & Menguc, B. (2005). The implications of socialisation and integration in supply chain management. *Journal of Operations Management*, *24,* pp. 604-620.

Crisp, R. J., & Turner, R. N. (2010). *Essential social psychology* (2nd ed.). London: Sage.

Day, G., & Nedungagi, P. (1994). Managerial representations of competitive advantage. *The Journal of Marketing, 58*(2), pp.31-44.

Dierickx, I., & Cool, K. (1989). Asset stock accumulation and sustainability of competitive advantage. *Management Science, 35*, pp. 1504-1511.

Dionisio, A., Menezes, R., & Mendes, D. A. (2005). *Uncertainty in financial markets, can entropy be a solution.* Lisboa: Universidade de Evora.

Dooley, K. (1997). A complex adaptive systems model of organisational change. *Non-Linear Dynamics, Phsycology and Life Sciences*, *1*(1), pp. 69-97.

Dorogovtsev, S. (2010). *Complex networks.* Oxford: Oxford University Press.

Easely, D., & Klienberg, J. (2010). *Networks, Crowds and Markets.* Cambridge: Cambridge University Press.

Easterby - Smith, M., Thorpe, R., & Jackson, P. (2012). Management research [Kindle Edition]. (4th). London: Sage. Retrieved from http://www.amazon.co.uk

Ebrahimi, N., Maasoumi, E., & Soofi, E. S. (1999). *Measuring informativeness of data by entropy and variance.* Retrieved 11 2011, from http://www.service.emory.edu/~mboerck/measureinform.pdf

Edwards, G., & Yearworth, M. (2011). Bridging from the phenomenological methods to systems interventions: a case study for SSM. *In Preparation for Submission to Systems Engineering* .

Efstathiou, J., Calinescu, A., & Blackburn, G. (2002). A web based expert system to assess the complexity of manufacturiing organisations. *Robotics and Computer Integrated Manufacturing*, *18*, pp. 305-311.

Efstathiou, J., Tassano, F., Sivadasan, S., Shirazi, R., Alvez, J., Frizelle, G., & Calinescu, A. (1999). Information complexity as a driver of emergent phenomena in the business community. In K. Ueda (Ed.), *Proceedings of the International Workshop on Emergent Synthesis* (pp. 1-6). Kobe.

Emmett, S., & Crocker, B. (2006). *The relationship driven supply chain* . Aldershot: Gower Publishing.

Erenguc, S. S., Simpson, N. C., & Vakharia, A. J. (1999). Integrated production distribution planning in supply chains: an invited review. *European Journal of Operations Research, 115*, pp. 219 - 236.

Feenberg, A. (2014). *The philosphy of praxis.* London: Verso.

Ferris, T. (2009). *On the methods of reserach for systems engineering.* Paper Presented at. 7th Annual Conference on Systems Engineering Research.

Frizelle, G. (1998). *The management of complexity in manufacturing.* London: Business Intelligence Ltd.

Frizelle, G., & Woodcock, E. (1995). Measuring complexity as an aid to developing operational strategy. *International Journal of Operations and Production Management, 15*(5), pp. 26-39.

Galbraith, J. R. (1974). Organisation design: an information processing review. *Interfaces, 4*(3), pp.28-36.

Gharajedaghi, J. (2011). *Systems thinking: managing chaos and complexity* (3rd ed.). London: Morgan Kaufmann.

Giannakis, M., Croom, S., & Slack, N. (2004). Supply chain paradigms. In S. New, & R. Westbrook (Eds.), *Understanding supply chains* (pp. 1-21). Oxford: Oxford University Press.

Gleik, J. (2011). *The Information - A history, a theory, a flood.* London: Fourth Estate.

Gonzales, L. S., Rubio, F. G., Gonzales, F. R., & Velthuis, M. P. (2010). Measurement and business processes: a systemic review. *Business Process Management Journal, 16*(1), pp. 114-134.

Grint, K. (2000). *The arts of leadership.* Oxford: Oxford University Press.

Gummesson, E. (2000). *Qualitative methods in management research* (2nd ed.). London: Sage.

Hamel, G. (2012). *What matter now: How to win in a world of relentless change, ferocious competition, and unstoppable innovation.* San Fransisco: Josey Bass.

Hannan, & Freeman. (2004). *Organisational Ecology.* London: McMillan.

Hannan, M. T., & Freeman, J. (1989). *Organisation ecology.* Cambridge: Harvard University Press.

Harrison, A., & van Hoek, R. (2008). *Logistics management & strategy* (3rd ed.). London: Pearson Education.

Hartley, R. V. (1927). Transmission of information. *Paper presented at the International congress of telegraphy and telephony.* Lake Como.

Hassard, J. (2001). Multiple paradigms and organisational analysis: a case study. *Organisational Studies, 12*(2), pp. 275-299.

Hayles, N. K. (1991). Complex dynamics in science and literature. In N. K. Hayles, *Chaos and order: complex dynamics in science and literature* (pp. 1-36). Chicago: University of Chicago Press.

Hussey, J., & Hussey, R. (1997). *Business research: a practical guide for undergraduate and postgraduate students.* Basingstoke: Palgrave.

Isik, F. (2010). An entropy based approach to measuring complexity in supply chains. *International Journal of Production Research, 48*(12).

Ivanov, D., & Sokolov, B. (2010). *Adaptive supply chain management.* London: Springer - Verlag.

Jackson, M. C., & Keys, P. (1984). Towards a systems of systems methodologies. *Journal of the Operational Research Society, 35*(6), pp 473 - 486. Retrieved Dec 28, 2012, from www.jstor.org

Jaynes, E. T. (1983). *Papers on probability, statistics and statistical physics.* (R. D. Rosenkrantz, Ed.) Dordrecht: Kluwer.

Jung, J. Y., Chin, C. H., & Cardoso, J. (2011). An entropy based uncertainty measure of process models. *Information Processing Letters, 111*, pp. 135-141.

Kapur, J. N., & Kesavan, H. K. (1992). *Entropy and optimisation principles with applications.* Academic Press.

Karberger, T., & Masoon, B. (2001). Entropy and economic process - physics perspective. *Ecological Economics, 36*, pp. 165-179.

Keynes, J. (1937). Alternative theories of the rate of interest. *The Economic Journal, 47*. Wiley. Retrieved 01 01, 2012, from http://www.jstor.org/stable/2225525

Kulza, K., & Nalepa, G. J. (2012). Proposal of square metrics for measuring business process complexity. *Proceedings of the federated conference on computer science and information systems*, (pp. 919 - 922). Warsaw.

Kumar, V. (1987). Entropic measures of manufacturing flexibility. *International Journal of Production Research, 25*(7), pp. 957-966.

Lassen, K. B., & van der Aalst, W. M. (2009). Complexity metrics with workflow nets. *Information and Software Technology*(51), pp. 610 - 626. doi:doi:10.1016/j.infsof.2008.08.005

Learmonth, M. (2008). Speaking out: Evidence based management: a backlash against pluralism in organisational studies. *Organisation, 15*(2), pp 283 - 291. doi:10.1177/1350508407087763

Lewin, K. (1946). Action reseach and minority problems. *Journal of Social Issues, 2*, pp. 34-36.

Li, M., & Vitanyi, P. (2008). *An introduction to kolmogorov complexity and its applications* (3rd ed.). New York: Springer.

Malpas, S., & Wake, P. (Eds.). (2013). *The routledge companion to critical and cultural theory* (2nd ed.). London, UK: Routledge.

Mendling, J., Reijers, H. A., & van der Aalst, W. (2010). Seven process modeling guidlines. *Information and Software Technology, 52*(2), pp.127 - 136. doi:doi:10.1016/j.infsof.2009.08.004

Merton, R. K. (1967). *On theoretical sociology*. New York: Free Press.

Milgate, M. (2001). Supply chain complexity and delivery performance: an international exploratory study. *Supply Chain Management: an International Journa*, *6*(3), pp. 106-118.

Mingers, J., & Rosenhead, J. (2004). Problem structuring methods in action. *European Journal of Operational Research, 152*, pp 530 - 554. doi:10.1016/S0377-2217(03)00056-0

Morgan, G. (2006). *Images of organisations.* London: Sage.

Newman, M. (2010). *Networks: an introduction.* Oxford: Oxford University Press.

North, M. J., & Macal, C. M. (2007). *Managing business complexity*. New York: Oxford University Press.

Prigogine, I. (1997). *The end of certainty*. New York: The Free Press.

Rao, H. A., & Gu, P. (1994). Flexibility in cellular manufacturing - an entropic measure. *The 10th ISPE International Conference on CAD/CAM Robotics and Factories of the Future*, (pp. 411-416).

Rassan, K. B., & van der Aalst, W. M. (2009). Complexity metrics for workflow nets. *Information and Software Technology, 51*(3), pp. 610-620.

Reza, M. (1994). *An introduction to information theory.* New York: Dover Publications.

Sanchez - Gonzales, L., Garcia, F., Ruiz, F., & Mendling, J. (2012). Quality indicators for business process models from a gateway complexity perspective. *Information and Software Technology*, pp 1159 - 1174. doi:10.1016/j.infsof.2012.05.001

Scuricini, G. B. (1988). Complexity in large technological systems. In L. Peliti, A. Vulpiani, L. Peliti, & A. Vulpiani (Eds.), *Measures of Complexity: Lecture Notes in Physics* (Vol. 314, pp. 83-101). Berlin: Springer-Verlag.

Shannon, C. E. (1948). A mathematical theory of communications. *Bell Systems Technical Journal, 27*, pp.379-423.

Shuiabi, E., Thomas, V., & Biuyan, N. (2005). Entropy as a measure of operational flexibility. *European Journal of Operational Reasearch, 165*, pp. 696-707.

Sivadasan, S., Efstathiou, J., Calinescu, A., & Huatuco, L. H. (2006). Advances on measuring the operational complexity of supplier - customer systems. *European Journal of Operational Research, 171*, pp208-226.

Sivadasan, S., Efstathiou, J., Frizelle, G., Shirazi, R., & Calinescu, A. (2002). An information theoretic methodolgy for measuring the operational complexity of customer supplier systems. *International Journal of Operations and Production Management, 22*(1), pp. 80-102.

Smith, J. A., Flowers, P., & Larkin, M. (2009). *Interpretive phenomenological analysis.* London: Sage.

Smolin, L. (2013). *Time reborn: from the crisis of physics to the future of the universe.* London: Allen Lane.

Sole, R. V., & Valverde, S. (2004). Information theory of complex networks:on evolution and architectural constraints. *Complex Networks, Lecture Notes in Physics, 650*, pp 189 - 207.

Soofi, E. S. (1994). Capturing the intangible concept of information. *Journal of the American Statistical Association, 89*(428), pp.1243-1254.

Soofi, E. S., & Gokhale, D. V. (1991). Minimum descrimination information estimator of the mean with known coefficient of variation. *Computational Statistics and Data Analysis, 11*, pp. 165-177.

Stacey, R. D. (2010). *Complexity & organisational reality* (2nd ed.). Oxford: Routledge.

Stebbins, R. A. (2006). Concatenated exploration: aiding theoretic memory by planning well for the future. *Journal of Contemporary Ethnography, 35*(5), pp. 483-494.

Sterman, J. D. (2000). *Business dynamic - systems thinking and modeling for a complex world.* New York: McGraw Hill.

Streatfield, P. (2001). *The paradox of control in organisations.* Abingdon: Routledge.

Tsallis, C. (2009). *Introduction to nonextensive statistical mechanics:approaching a complex world.* New York: Springer.

Vanderfesten, I. T., & Riejers, H. A. (2004). *Cohesion and coupling metrics for workflow process design.* Eindhoven: Einhoven University of Technology.

Wang, Y., Potter, A., & Naim, M. (1997). Electronic marketplaces for tailored logistics. *Industrial Management and Data Systems, 107*(8), pp.1170 - 1187.

Welch, C., & Wilkinson, I. (2002). Idean logic and network theory in business marketting. *Journal of Business to Business Marketing, 9*(3), pp.27 - 48. Retrieved 08 24, 2010, from http://www.informaworld.com/smpp/title~content=t792303971

Whalley, P. B. (2004). *Basic Engineering Thermodynamics.* Oxford: Oxford University Press.

Wilding, R. (1988). The supply chain complexity: uncertainty generation in the supply chain. *International Journal of Physical Distribution & Logistics Management, 28*(8), pp 599-616.

Williamson, O. (1975). *Markets and hierarchies: analysis and antitrust implications.* New York: Free Press.

Yao, D. D. (1985). Material information flows in flexible manufacturing systems. *Material Flow, 2*, pp. 143-149.

Appendix

Appendix A: Proofs of the form H

Two proofs are provided. Firstly that provided by Ben-Naim (2012); secondly, that provided by Shannon (1948). Alternative proofs are available (Ben-Naim himself refers to Khinchin (1957) and Katz (1967); however, these proofs are useful because they add detail to the arguments proposed in the thesis.

Ben-Naims proof of the form H

Firstly, the assumption is made, following Shannon (1948), that:

a) H is continuous in all $p_{i.}$

b) If all p_i are equal ($p_i = 1/n$) then H should have a maximum value which is a monotonically increasing function of n.

c) If choice is broken down into successive choices, the quantity H should be the weighted sum of the individual values.

An experiment has n outcomes, $A_1, \ldots, A_n$. The outcomes are grouped in to r sets, each of which contains m_k elements such that $\sum_{K=1}^{r} m_k = n$. $A'_1, \ldots A'_r$ are the new events defined in terms of the original outcomes:

$$A'_1 = \{A_1 + A_2 + \cdots + A_{m1}\}$$

$$A'_2 = \{A_{m1+1} + A_{m1+2} + \cdots + A_{m1+m2}\}$$

$$A'_3 = \{A_{m1+m2+1} + A_{m1+m2+2} + \cdots + A_{m1+m2+m3}\}$$

$$\ldots$$

$$A'_r = \{A_{\sum_{k=1}^{r-1} m_k+1} + \cdots + A_{n=\sum_{k=1}^{r} m_k}\}$$

Assuming the original events to be mutually exclusive, the probabilities of the new events are the sum of the probabilities of the original events included in the new event.

$$p'_1 = P(A'_1) = \sum_{i=1} p_i$$

$$p'_2 = P(A'_2) = \sum_{i=m_1+1}^{m_1+m_2} p_i$$

$$\ldots$$

$$p'_r = P(A'_r) = \sum_{i=m_k+1}^{m_1+m_2} p_i$$

This structure can be represented as a table, see Table 73:

Table 73: Relationship between Outcomes and Groupings.

	n Outcomes				
R Groups		x_1	x_2	x_3	x_n
	m^1	$k_{x_1}^{m_1}$			
	m^2				
	m^3				
	m^r				$k_{x_n}^{m_r}$

Table 74: Relationship between Outome Probabilities and Grouping Probabilities.

	n Outcome Probabilities				
r Group Probabilities		x_1	x_2	x_3	x_n
	p^{m_1}	$p_{x_1}^{k_1}$			
	p^{m_2}				
	p^{m_3}				
	p^{m_r}				$p_{x_n}^{k_r}$

To be consistent with the three assumptions above H from the original n outcomes must be equal to H from the r groupings plus and average H within the group events.

$$H(p_1 \dots p_n) = H(p'_1, \dots, p'_r) + \sum_{k=1}^{r} p'_k \, H\left(\frac{p_1^k}{p'_k}, \dots, \frac{p_{m_k}^k}{p'_k}\right) \qquad \text{Equation 80}$$

The next step is to 'normalise' the probabilities of each outcome by creating a third level of events. This time the new events normalise the outcomes of each of the original events by denoting the original events $A_1, \dots, A_n$ as consisting of M_1 elements, all of equal probability. For example, If $p_i = 0.04$ then $p_i = 4M$ where $M = 0.01$.

Modifying Equation 80 for the equiprobable states we get:

$$H\left(\frac{1}{M}, \dots, \frac{1}{M}\right) = H(p_1, \dots, p_n) + \sum_{i=1}^{n} p_i \, H\left(\frac{1/M}{M_i/M}, \dots, \frac{1/M}{M_i/M}\right) \qquad \text{Equation 81}$$

The function $F(m)$ is defined as

$$F(M) = H\left(\frac{1}{M}, \dots, \frac{1}{M}\right)$$

and rewrite Equation 81 as

$$F(M) = H(p_1, \dots, p_n) + \sum_{i=1}^{n} p_i \, \mathrm{F}(M_i) \qquad \text{Equation 82}$$

Ben-Naim then digresses to find the function $F(m)$. To do so he chooses a particular state where all m are equal, such that:

$$p_i = \frac{M_i}{M} = \frac{m}{M} = \frac{1}{n}.$$

Therefore, in this case:

$$F(n) = H(p_1, \dots, p_n) = H\left(\frac{1}{n}, \dots, \frac{1}{n}\right) \qquad \text{Equation 83}$$

and

$$\sum_{i=1}^{n} p_i \, F(M_i) = F(m).$$

Consequently, Equation 82 can be reduced to:

$$F(n \times m) = F(m) + F(n)$$

and the only function that has this property is the logarithm function, and

$$\log(n \times m) = \log(m) + \log(n).$$

Without being specific on the base of the logarithm,

$$F(M) = \log M.$$

Having found the form of the function $F(M)$, Ben-Naim returns to Equation 84

$$H(p_1, \dots, p_n) = F(M) + \sum_{i=1}^{n} p_i \, \mathrm{F}(M_i)$$ Equation 84

$$H(p_1, \dots p_n) = \log M - \sum_{x=1}^{n} p_i \log M_i$$ Equation 85

$$H(p_1, \dots p_n) = -\sum_{i=1}^{n} p_i \log \frac{M_i}{M}$$ Equation 86

$$H(p_1, \dots, p_n) = -\sum_{i=1}^{n} p_i \log p_i$$ Equation 87

With Equation 87 being the general form of the function H.

Shannon's proof of the form H

Let

$$H\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right) = A(n).$$

$A(n)$ can be considered a choice from S^m equally likely possibilities into a series of m choices from s equally likely possibilities and obtain

$$A(S^m) = mA(s).$$

Similarly

$$A(t^n) = nA(t).$$

We can chose *n* arbitrarily large and find an *m* to satisfy

$$s^m \leq t^n < s^{m+1}.$$

Thus, taking the logarithm and dividing by $n \, logs,$

$$\frac{m}{n} \leq \frac{\log t}{\log s} \leq \frac{m}{n} + \frac{1}{n} \quad or \quad \left|\frac{m}{n} - \frac{\log t}{\log s}\right| < \epsilon$$

Where ϵ is arbitrarily small. Now from the monotonic property of $A(n)$,

$$A(s^m) \leq A(t^n) \leq A(s^{m+1})$$

$$mA(s) \leq nA(t) \leq (m+1)A(s).$$

Hence, dividing by $nA(s)$

$$\frac{m}{n} \leq \frac{A(t)}{A(s)} \leq \frac{m}{n} + \frac{1}{n} \; or \; \left|\frac{m}{n} - \frac{A(t)}{A(s)}\right| < \epsilon$$

$$\left|\frac{A(t)}{A(s)} - \frac{\log t}{\log s}\right| < 2\epsilon \quad A(t) = K \log t$$ Equation 88

where K is positive.

Shannon then continues: Suppose there is a choice from *n* possibilities with commensurate probabilities $P_i = \frac{n_i}{\sum n_i}$ where n_i are integers. Choice can be broken down from $\sum n_i$ possibilities into a choice from *n* possibilities with probabilities $p_i, \dots, p_n$ and then, if the *i*th was chosen, a choice from n_i with equal probabilities. This can be equated to total choice from $\sum n_i$ by two methods:

$$K \log \sum n_i = H(p_1, \dots, p_n) + K \sum P_i \log n_i.$$

Hence

$$H = K\left[\sum p_i Log \sum n_i - \sum p_i log\, n_i\right] = -K \sum p_i log \frac{n_i}{\sum n_1} = -K \sum p_i log p_i$$

which is a general form of the function H.

Appendix B: The construction of the adjacency matrix from the SCOR model.

For this research the Supply Chain Operating Reference Model (version 10) has been used as a structure from which an adjacency matrix was constructed. This appendix explains the model in more detail and explains how the adjacency matrix is constructed. Appendix C contains the tables for each of the plan, source, make, deliver and return subcomponents in the model.

The supply chain operating reference model presents a structure for defining a model for supply chain operations. The model is broken down into five major subsystems: Plan, Source, Make, Deliver and Return. Each of these subsystems is further deconstructed into two layers. For example, the top level of the model is the SCOR model, this deconstructs to the five level one subsystems, plan, source, make, deliver and return. Each subsystem further deconstructs into an operational and monitoring component and each of these deconstructs into a number of level three processes. An example of the process breakdown is shown in Figure 56. Appendix C contains tables for each of the level one processes, each detailing the level two and level three processes.

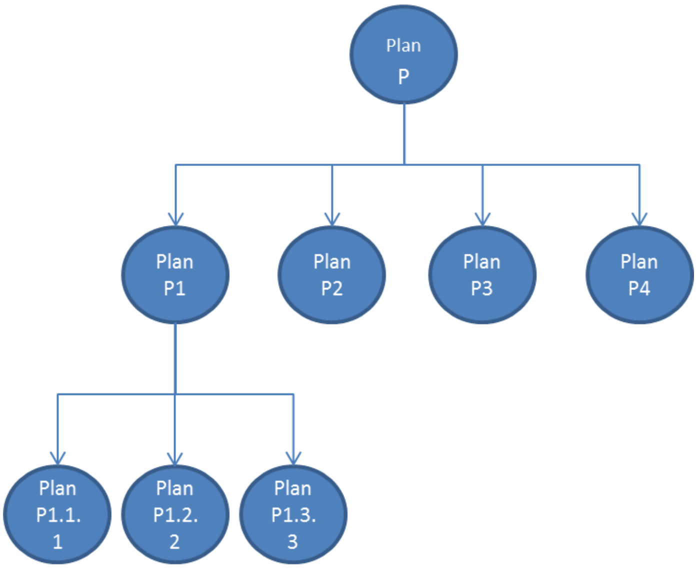


Figure 56: Example of the process hierarchy for the SCOR model

The adjacency matrix is constructed by creating a spreadsheet with the level three, two, and one process constructs on a horizontal and vertical axes. Figure 29 in section six shows this construct for level two and three processes for the plan process set. Where a process link exists, the value one is added to the adjacency matrix. Most of the SCOR model contains uni-directional links, thus the value one is added at the intersect where the link leaves the process on the vertical axis and enters the process on the horizontal axis. Where the link is bi-directional a value one is added to both the vertical out/horizontal in, and the horizontal out/vertical in; in this case the values are symmetrical about the diagonal.

In some cases a link exists that connects to processes outside the supply chain operating reference model; these connections may be incoming or outgoing. To capture these links an additional category is added to the adjacency matrix, non-graph in (NGI) for the in degree, and non-graph out (NGO) for the out degree. This thesis uses out degree as an independent variable; consequently the non-graph out category is included in the analysis.

Appendix C: Tables for the supply chain operating reference model

Tables for the Planning processes

sP1	sP1.1	Identify, Prioritise and Aggregate Supply Chain Requirements
	sP1.2	Identify, Prioritise and Aggregate Supply Chain Resources
	sP1.3	Balance Supply Chain Resources with Supply Chain Requirements
	sP1.4	Establish and Communicate Supply Chain Plans
sP2	sP2.1	Identify, Prioritise and Aggregate Product Requirements
	sP2.2	Identify, Assess and Aggregate Product Resources
	sP2.3	Balance Product Resources with Product Requirements
	sP2.4	Establish Sourcing Plans
sP3	sP3.1	Identify, Prioritize and Aggregate Production Requirements
	sP3.2	Identify Assess and Aggregate Production Resources
	sP3.3	Balance Production Resources with Production Requirements
	sP3.4	Establish Production Plans
ssP4	sP4.1	Identify, Prioritize and Aggregate Delivery Requirements
	sP4.2	Identify, Assess and Aggregate delivery Resources
	sP4.3	Balance Delivery Resources and Capabilities with delivery Requirements
	sP4.4	Establish Delivery Plans
sP5	sP5.1	Assess and Aggregate Return Requirements
	sP5.2	Identify, Assess and Aggregate Return Resources
	sP5.3	Balance Return Resources with Return Requirements
	sP5.4	Establish and Communicate Return Plans

sEP1	sEP1	Manage Business Rules for Plan Process
sEP2	sEP2	Manage Performance of Supply Chain
sEP3	sEP3	Manage Plan Data Collection
sEP4	sEP4	Manage Integrated Supply Chain Inventory

sEP5	sEP5	Manage Integrated Supply Chain Capital Assets
sEP6	sEP6	Manage Integrated Supply Chain Transportation
sEP7	sEP7	Manage Planning Configuration
sEP8	sEP8	Manage Plan Regulatory Requirements and Compliance
sEP9	sEP9	Manage Supply Chain Plan Risks
sEP10	sEP10	Align Supply Chain Unit Plan with Financial Plan

Tables for the Sources processes.

sS1	sS1.1	Schedule Product Deliveries
	sS1.2	Receive Product
	sS1.3	Verify Product
	sS1.4	Transfer Product
	sS1.5	Authorise Supplier Payments
sS2	sS2.1	Schedule Product Deliveries
	sS2.2	Process Element: Receive Product
	sS2.3	Verify Product
	sS2.4	Transfer Product
	sS2.5	Authorise Supplier Payments
sS3	sS3.1	Identify Sources of Supply
	sS3.2	Select Final Supplier(s) and Negotiate
	sS3.3	Schedule Product Deliveries
	sS3.4	Receive Product
	sS3.5	Verify Product
	sS3.6	Transfer Product
	sS3.7	Authorise Supplier Payments

sES1	sES1	Manage Sourcing Business Rules
sES2	sES2	Assess Supplier Performance
sES3	sES3	Maintain Source Data
sES4	sES4	Manage Product Inventory
sES5	sES5	Manage Capital Assets
sES6	sES6	Manage Incoming Product

sES7	sES7	Manage Supplier Network
sES8	sES8	Manage Import/Export Requirements
sES9	sES9	Manage supply Chain Source Risk
sES10	sES10	Manage Supplier Agreements

Tables for the Make process

sM1	sM1.1	Schedule Production Activities
	sM1.2	Issue Material
	sM1.3	Produce and Test
	sM1.4	Package
	sM1.5	Stage Product
	sM1.6	Release Product to Delivery
	sM1.7	Waste Disposal
sM2	sM2.1	Schedule Production Activities
	sM2.2	Issue Sourced/In Process Product
	sM2.3	Produce and Test
	sM2.4	Package
	sM2.5	Stage Finished Product
	sM2.6	Release Finished Product to Delivery
	sM2.7	Waste Disposal
sM3	sM3.1	Finalise Production Engineering
	sM3.2	Schedule Production Activities
	sM3.3	Issue Sources/In Process Product
	sM3.4	Produce and Test
	sM3.5	Package
	sM3.6	Stage Finished Product
	sM3.7	Release Product to Delivery
	sM3.8	Waste Disposal

sEM1	sEM1	Manage Production Rules
sEM2	sEM2	Manage Production Performance
sEM3	sEM3	Manage Make Information
sEM4	sEM4	Manage In Process Products (WIP)

sEM5	sEM5	Manage Make Equipment and Facilities
sEM6	sEM6	Manage Transportation (WIP)
sEM7	sEM7	Manage Production Network
sEM8	sEM8	Manage Make Regulatory Environment
sEM9	sEM9	Manage Supply Chain Make Risk

Tables for the Deliver processes

sD1	sD1.1	Process Enquiry and Quote
	sD1.2	Receive, Enter and Validate Order
	sD1.3	Reserve Inventory and Determine Delivery Date
	sD1.4	Consolidate Orders
	sD1.5	Build Loads
	sD1.6	Route Shipment
	sD1.7	Select Carriers and Rate Shipment
	sD1.8	Receive Product from Source or Make
	sD1.9	Pick Product
	sD1.10	Pack Product
	sD1.11	Load Vehicle and Generate Shipping Docs
	sD1.12	Ship Product
	sD1.13	Receive and Verify Product by Customer
	sD1.14	Install Product
	sD1.15	Invoice
sD2	sD2.1	Process Inquiry and Quote
	sD2.2	Receive, Configure, Enter and Validate Order
	sD2.3	Reserve Inventory and Determine Delivery Date
	sD2.4	Consolidate Order
	sD2.5	Build Loads
	sD2.6	Route Shipment
	sD2.7	Select Carriers and Rate Shipment
	sD2.8	Receive Product from Source or Make
	sD2.9	Pick Product
	sD2.10	Pack Product
	sD2.11	Load Product and generate Shipping Docs
	sD2.12	Ship Product
	sD2.13	Receive and Verify Product by Customer
	sD1.14	Install Product
	sD2.15	Invoice
sD3	sD3.1	Obtain and Respond to RFP/RFQ

	sD3.2	Negotiate and Receive Contract
	sD3.3	Enter Order, Commit Resources and Launch Program
	sD3.4	Schedule Installation
	sD3.5	Build Loads
	sD3.6	Route Shipments
	sD3.7	Select Carrier and Rate Shipments
	sD3.8	Receive Product from Source or Make
	sD3.9	Pick Product
	sD3.10	Pack Product
	sD3.11	Load Product and Generate Shipping Docs
	sD3.12	Ship Product
	sD3.13	Receive and verify Product by Customer
	sD3.14	Install Product
	sD3.15	Invoice
sD4	sD4.1	Generate Stocking Schedule
	sD4.2	Receive Product at the Store
	sD4.3	Pick Product from Backroom
	sD4.4	Stock Shelf
	sD4.5	Fill Shopping Cart
	sD4.6	Checkout
	sD4.7	Deliver and/or Install

sED1	sED1	Manage Delivery Business Rules
sED2	sED2	Assess delivery Performance
sED3	sED3	Manage Delivery Information
sED4	sED4	Manage Finished Goods Inventories
sED5	sED5	Manage Delivery capital Assets
sED6	sED6	Manage Transportation

sED7	sED7	Manage Product Life Cycle
sED8	sED8	Manage Import/Export Requirements
sED9	sED9	Manage Supply Chain Delivery Risks

Tables for the Return processes

sSR1	sSR1.1	Identify Defective Product Condition
	sSR1.2	Disposition Defective Product
	sSR1.3	Request Defective Product Return Authorisation
	sSR1.4	Schedule Defective Product Shipment
	sSR1.5	Return Defective Product
sDR1	sDR1.1	Authorise Defective Product Return
	sDR1.2	Schedule Defective Return Receipt
	sDR1.3	Receive Defective Product
	sDR1.4	Transfer Defective Product
sSR2	sSR2.1	Identify MRO Product Condition
	sSR2.2	Disposition MRO Product
	sSR2.3	Request MRO return Authorisation
	sSR2.4	Schedule MRO Shipment
	sSR2.5	Return MRO Product
sDR2	sDR2.1	Authorise MRO Product Return
	sDR2.2	Schedule MRO Return Receipt
	sDR2.3	Receive MRO Product

	sDR2.4	Transfer MRO Product
sSR3	sSR3.1	Identify Excess Product Condition
	sSR3.2	Disposition Excess Product
	sSR3.3	Request Excess Product Return Authorisation
	sSR3.4	Schedule Excess Product Shipment
	sSR3.5	Return Excess Product
sDR3	sDR3.1	Authorise Excess Product Return
	sDR3.2	Schedule Excess Return Receipt
	sDR3.3	Receive Excess Product
	sDR3.4	Transfer Excess Product

sER1	sER1	Manage Business Rules for Return Process
sER2	sER2	Manage Performance of Return Processes
sER3	sER3	Manage Return Data Collection
sER4	sER4	Manage Return Inventory
sER5	sER5	Manage Return Capital Assets
sER6	sER6	Manage Return Transportation
sER7	sER7	Manage Return Network Configuration

sER8	sER8	Manage Return Regulatory Requirements and Compliance
sER9	sER9	Manage Supply Chain Return Risks

Appendix D

The process for explaining and completing the questionnaire presented to the supply chain subject matter experts consisted of an explanation and a set of questions. For ease of completion the questions were made available at www.surveymonkey.com/s/3HNGFNH); all the respondents choose to complete the online questionnaire. The geographic location of the respondents – UK, Germany, Bahrain and Saudi Arabia – meant that all, except one, interviews were conducted over the internet using skype or Microsoft lync. The next section sets out the process used in the interview and the section following details the questionnaire.

Process Steps

Process and questions for the structured interviews used in the field research:

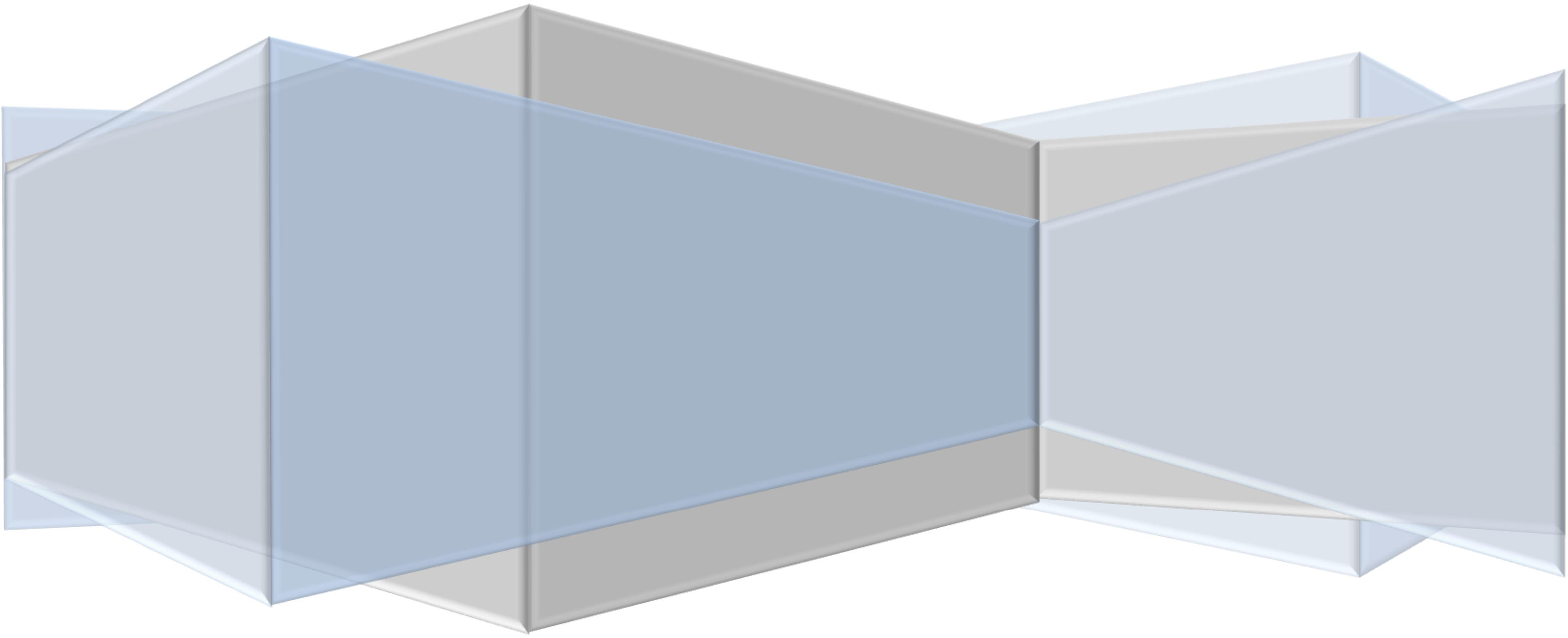
Process Step	Explanation/Question
1. Introduction	Explain the purpose of the interview: Thank you for taking part. The purpose of this interview is to get your expert view on the sets of supply chain processes necessary to operate one of two different business scenarios: The scenario where an organisation acts as a prime-contractor, and the scenario where the organisation acts as a sub-contractor. The selection process will use the supply chain operating reference model.
2. Checkpoint	From your 'operational' perspective, will you be able to relate business scenarios to the SCOR framework? Yes/No. If yes, continue; if no: Thank you, It is important I collect data only from experts who are comfortable with the application of the SCOR model. If you feel you will not be comfortable relating

Process Step	Explanation/Question
	business scenarios to the SCOR model then we will conclude the interview. Thank you for your time.
3. Scenario Build	Can you please take a few minutes to consider either a prime or a sub-contractor scenario, ideally one that you use in your operation.
4. Questionnaire	Using your business scenarios can you now complete the attached/online questionnaire which will ask you to select processes you consider to be within your scenario. The online version will guide you through the SCOR model depending on your answers. If filling a paper version, please note follow the process carefully. Please note that your responses will be anonymous.
5. Thank you	Afterwards: Thanks for taking the time to complete the questionnaire. As I said, the selections you have entered will be anonymous. These will now be collated and used to inform and validate the data I will use in my analysis. If you like, once complete, I would be happy to arrange to take you through the major findings of my research.|I will contact you again once it is complete.

Questionnaire

The selection of process scopes for the Prime and Sub-Contractor business scenarios

Peter Dickinson



Questionnaire

The following pages detail the plan, source, make, deliver and return processes in the SCOR framework. To ensure the data is captured correctly, firstly, please indicate if this business scenario is a prime or a sub-contractor set of the framework.

	Please tick the appropriate box
This scenario represents a sub-contractor set of the framework	☐
This scenario represents a prime contractor of the framework	☐

If you have chosen the sub-contractor set of the framework, please tick all the processes that will be required.

If you have chosen the prime contractor set of the framework, please tick the processes that will be required IN ADDITION to the full framework.

Plan Processes

If you believe all the planning processes will be required for this scenario, please check the box below.

All planning processes are required	☐

If you have checked the above box, please move to the SOURCE section.

If not, please check the box identifying the level two or level three processes required. If you tick a level two process, there is no need to tick the relevant level three processes.

Level 2		Level 3		
sP1		sP1.1		Identify, Prioritise and Aggregate Supply Chain Requirements
		sP1.2		Identify, Prioritise and Aggregate Supply Chain Resources
		sP1.3		Balance Supply Chain Resources with Supply Chain Requirements
		sP1.4		Establish and Communicate Supply Chain Plans
sP2		sP2.1		Identify, Prioritise and Aggregate Product Requirements
		sP2.2		Identify, Assess and Aggregate Product Resources
		sP2.3		Balance Product Resources with Product Requirements
		sP2.4		Establish Sourcing Plans
sP3		sP3.1		Identify, Prioritize and Aggregate Production Requirements
		sP3.2		Identify Assess and Aggregate Production Resources
		sP3.3		Balance Production Resources with Production Requirements

Level 2		Level 3		
		sP3.4		Establish Production Plans
sP4		sP4.1		Identify, Prioritize and Aggregate Delivery Requirements
		sP4.2		Identify, Assess and Aggregate delivery Resources
		sP4.3		Balance Delivery Resources and Capabilities with delivery Requirements
		sP4.4		Establish Delivery Plans
sP5		sP5.1		Assess and Aggregate Return Requirements
		sP5.2		Identify, Assess and Aggregate Return Resources
		sP5.3		Balance Return Resources with Return Requirements
		sP5.4		Establish and Communicate Return Plans

sEP1	☐	Manage Business Rules for Plan Process
sEP2	☐	Manage Performance of Supply Chain

sEP3	☐	Manage Plan Data Collection
sEP4	☐	Manage Integrated Supply Chain Inventory
sEP5	☐	Manage Integrated Supply Chain Capital Assets
sEP6	☐	Manage Integrated Supply Chain Transportation
sEP7	☐	Manage Planning Configuration
sEP8	☐	Manage Plan Regulatory Requirements and Compliance
sEP9	☐	Manage Supply Chain Plan Risks
sEP10	☐	Align Supply Chain Unit Plan with Financial Plan

Source processes.

If you believe all the source processes will be required for this scenario, please check the box below.

All source processes are required	☐

If you have checked the above box, please move to the MAKE section.

If not, please check the box identifying the level two or level three processes required. If you tick a level two process, there is no need to tick the relevant level three processes.

Level 2		Level 3		
sS1	☐	sS1.1	☐	Schedule Product Deliveries
		sS1.2	☐	Receive Product
		sS1.3	☐	Verify Product
		sS1.4	☐	Transfer Product
		sS1.5	☐	Authorise Supplier Payments
sS2	☐	sS2.1	☐	Schedule Product Deliveries
		sS2.2	☐	Process Element: Receive Product

Level 2		Level 3		
		sS2.3	☐	Verify Product
		sS2.4	☐	Transfer Product
		sS2.5	☐	Authorise Supplier Payments
sS3	☐	sS3.1	☐	Identify Sources of Supply
		sS3.2	☐	Select Final Supplier(s) and Negotiate
		sS3.3	☐	Schedule Product Deliveries
		sS3.4	☐	Receive Product
		sS3.5	☐	Verify Product
		sS3.6	☐	Transfer Product
		sS3.7	☐	Authorise Supplier Payments

sES1	☐	Manage Sourcing Business Rules

sES2	☐	Assess Supplier Performance
sES3	☐	Maintain Source Data
sES4	☐	Manage Product Inventory
sES5	☐	Manage Capital Assets
sES6	☐	Manage Incoming Product
sES7	☐	Manage Supplier Network
sES8	☐	Manage Import/Export Requirements
sES9	☐	Manage supply Chain Source Risk
sES10	☐	Manage Supplier Agreements

Make Process

If you believe all the make processes will be required for this scenario, please check the box below.

All make processes are required	☐

If you have checked the above box, please move to the DELIVER section.

If not, please check the box identifying the level two or level three processes required. If you tick a level two process, there is no need to tick the relevant level three processes.

Level 2		Level 3		
sM1	☐	sM1.1	☐	Schedule Production Activities
		sM1.2	☐	Issue Material
		sM1.3	☐	Produce and Test
		sM1.4	☐	Package
		sM1.5	☐	Stage Product
		sM1.6	☐	Release Product to Delivery
		sM1.7	☐	Waste Disposal

Level 2		Level 3		
sM2	☐	sM2.1	☐	Schedule Production Activities
		sM2.2	☐	Issue Sourced/In Process Product
		sM2.3	☐	Produce and Test
		sM2.4	☐	Package
		sM2.5	☐	Stage Finished Product
		sM2.6	☐	Release Finished Product to Delivery
		sM2.7	☐	Waste Disposal
sM3	☐	sM3.1	☐	Finalise Production Engineering
		sM3.2	☐	Schedule Production Activities
		sM3.3	☐	Issue Sources/In Process Product
		sM3.4	☐	Produce and Test
		sM3.5	☐	Package

Level 2		Level 3		
		sM3.6	☐	Stage Finished Product
		sM3.7	☐	Release Product to Delivery
		sM3.8	☐	Waste Disposal

sEM1	☐	Manage Production Rules
sEM2	☐	Manage Production Performance
sEM3	☐	Manage Make Information
sEM4	☐	Manage In Process Products (WIP)
sEM5	☐	Manage Make Equipment and Facilities
sEM6	☐	Manage Transportation (WIP)
sEM7	☐	Manage Production Network
sEM8	☐	Manage Make Regulatory Environment
sEM9	☐	Manage Supply Chain Make Risk

Deliver Processes

If you believe all the deliver processes will be required for this scenario, please check the box below.

All deliver processes are required	☐

If you have checked the above box, please move to the RETURN section.

If not, please check the box identifying the level two or level three processes required. If you tick a level two process, there is no need to tick the relevant level three processes.

Level 2		Level 3		
sD1	☐	sD1.1	☐	Process Enquiry and Quote
		sD1.2	☐	Receive, Enter and Validate Order
		sD1.3	☐	Reserve Inventory and Determine Delivery Date
		sD1.4	☐	Consolidate Orders
		sD1.5	☐	Build Loads
		sD1.6	☐	Route Shipment

Level 2		Level 3		
		sD1.7	☐	Select Carriers and Rate Shipment
		sD1.8	☐	Receive Product from Source or Make
		sD1.9	☐	Pick Product
		sD1.10	☐	Pack Product
		sD1.11	☐	Load Vehicle and Generate Shipping Docs
		sD1.12	☐	Ship Product
		sD1.13	☐	Receive and Verify Product by Customer
		sD1.14	☐	Install Product
		sD1.15	☐	Invoice
sD2	☐	sD2.1	☐	Process Inquiry and Quote
		sD2.2	☐	Receive, Configure, Enter and Validate Order
		sD2.3	☐	Reserve Inventory and Determine Delivery Date

Level 2		Level 3		
		sD2.4	☐	Consolidate Order
		sD2.5	☐	Build Loads
		sD2.6	☐	Route Shipment
		sD2.7	☐	Select Carriers and Rate Shipment
		sD2.8	☐	Receive Product from Source or Make
		sD2.9	☐	Pick Product
		sD2.10	☐	Pack Product
		sD2.11	☐	Load Product and generate Shipping Docs
		sD2.12	☐	Ship Product
		sD2.13	☐	Receive and Verify Product by Customer
		sD1.14	☐	Install Product
		sD2.15	☐	Invoice

Level 2		Level 3		
sD3	☐	sD3.1	☐	Obtain and Respond to RFP/RFQ
		sD3.2	☐	Negotiate and Receive Contract
		sD3.3	☐	Enter Order, Commit Resources and Launch Program
		sD3.4	☐	Schedule Installation
		sD3.5	☐	Build Loads
		sD3.6	☐	Route Shipments
		sD3.7	☐	Select Carrier and Rate Shipments
		sD3.8	☐	Receive Product from Source or Make
		sD3.9	☐	Pick Product
		sD3.10	☐	Pack Product
		sD3.11	☐	Load Product and Generate Shipping Docs
		sD3.12	☐	Ship Product

Level 2		Level 3		
		sD3.13	☐	Receive and verify Product by Customer
		sD3.14	☐	Install Product
		sD3.15	☐	Invoice
sD4	☐	sD4.1	☐	Generate Stocking Schedule
		sD4.2	☐	Receive Product at the Store
		sD4.3	☐	Pick Product from Backroom
		sD4.4	☐	Stock Shelf
		sD4.5	☐	Fill Shopping Cart
		sD4.6	☐	Checkout
		sD4.7	☐	Deliver and/or Install

sED1	☐	Manage Delivery Business Rules
sED2	☐	Assess delivery Performance
sED3	☐	Manage Delivery Information
sED4	☐	Manage Finished Goods Inventories
sED5	☐	Manage Delivery capital Assets
sED6	☐	Manage Transportation
sED7	☐	Manage Product Life Cycle
sED8	☐	Manage Import/Export Requirements
sED9	☐	Manage Supply Chain Delivery Risks

Return processes

If you believe all the return processes will be required for this scenario, please check the box below.

All return processes are required	☐

If you have checked the above box, you have completed the questionnaire. Thank you for your time.

If not, please check the box identifying the level two or level three processes required. If you tick a level two process, there is no need to tick the relevant level three processes.

Level 2		Level 3		
sSR1	☐	sSR1.1	☐	Identify Defective Product Condition
		sSR1.2	☐	Disposition Defective Product
		sSR1.3	☐	Request Defective Product Return Authorisation
		sSR1.4	☐	Schedule Defective Product Shipment
		sSR1.5	☐	Return Defective Product
sDR1	☐	sDR1.1	☐	Authorise Defective Product Return

Level 2		Level 3		
		sDR1.2	☐	Schedule Defective Return Receipt
		sDR1.3	☐	Receive Defective Product
		sDR1.4	☐	Transfer Defective Product
sSR2	☐	sSR2.1	☐	Identify MRO Product Condition
		sSR2.2	☐	Disposition MRO Product
		sSR2.3	☐	Request MRO return Authorisation
		sSR2.4	☐	Schedule MRO Shipment
		sSR2.5	☐	Return MRO Product
sDR2	☐	sDR2.1	☐	Authorise MRO Product Return
		sDR2.2	☐	Schedule MRO Return Receipt
		sDR2.3	☐	Receive MRO Product
		sDR2.4	☐	Transfer MRO Product

Level 2		Level 3		
sSR3	☐	sSR3.1	☐	Identify Excess Product Condition
		sSR3.2	☐	Disposition Excess Product
		sSR3.3	☐	Request Excess Product Return Authorisation
		sSR3.4	☐	Schedule Excess Product Shipment
		sSR3.5	☐	Return Excess Product
sDR3	☐	sDR3.1	☐	Authorise Excess Product Return
		sDR3.2	☐	Schedule Excess Return Receipt
		sDR3.3	☐	Receive Excess Product
		sDR3.4	☐	Transfer Excess Product

sER1	☐	Manage Business Rules for Return Process
sER2	☐	Manage Performance of Return Processes

sER3	☐	Manage Return Data Collection
sER4	☐	Manage Return Inventory
sER5	☐	Manage Return Capital Assets
sER6	☐	Manage Return Transportation
sER7	☐	Manage Return Network Configuration
sER8	☐	Manage Return Regulatory Requirements and Compliance
sER9	☐	Manage Supply Chain Return Risks

You have completed the questionnaire.

Thank you very much for your time.

Appendix E

The Pattern for the development of qpalzm56tygv

The pattern for describing the string ‘qpalzm56tygv’ can be written as follows: On a standard desktop computer keyboard beginning on the left, select the leftmost and then the right most letter from the top row of letters. Repeat this process for the second and third row. Using the same keyboard and including the row of numerals, starting with the row of numerals, where an even number of characters exist, select the middle two characters. Where an odd number exists, select the only the middle character. Repeat for the three rows of letters starting with the top row. This describe the pattern that develops the string ‘qpalzm56tygv’.

It is possible that with some refinement the description may be marginally reduced in length; however, in its present form it consists of eighty five words and four hundred and ninety one characters (including spaces as delineating characters). Obviously this is longer than the string itself, and given the significant difference in size between the pattern description and the string itself, it is very likely that, in this case, the shortest description is the string itself.