
The descriptive properties of prescriptive theories: an application of systems thinking in data warehousing

ROELIEN GOEDE¹

Abstract

Information systems and in particular data warehouses are very expensive systems to develop. It is therefore not advisable to experiment with ideas too different from current practices. This makes it difficult to apply prescriptive theories in an existing field. From theoretical considerations one might want to develop a data warehouse according to another method such as critical systems thinking methodology. It is however very difficult to persuade data warehouse practitioners to attempt such an experiment. This might be because they would rather adhere to known practices or that they are not sufficiently knowledgeable on critical systems thinking (or any other prescriptive theory) to apply it to such an expensive project. This paper describes a method in which prescriptive theories may be used descriptively to analyse their applicability in a specific field of application. The proposed method is used to understand the practices of the data warehouse discipline from the perspectives of the systems thinking discipline. It is also indicated how this method could be used in other studies where the behaviour of participants is viewed from a point of view of which the detail are unknown to the participants.

Keywords: Data warehousing, Systems thinking, Prescriptive theory, Descriptive theory, Interpretative research.

Disciplines: Information technology, systems theory, data warehousing, hermeneutics,

Introduction

This paper explores the nature of theory in terms of descriptive and prescriptive theories. Descriptive theories are mostly used to describe events that took place in the past. Prescriptive theory is used to guide future action. The paper explores the possibility to use prescriptive theories descriptively as well as the benefits of doing so. A method for using prescriptive theories descriptively is proposed and applied in using systems thinking methods in data warehouse development.

The paper starts with a short discussion on the descriptive and prescriptive nature of theories. The main part of this paper presents a method to explore the descriptive value of prescriptive theories. This discussion begins in section 3 with a short description of systems thinking

¹. Department of Computer Science and Information Systems, North-West University, Vaal Triangle Campus, South Africa, E-mail: roelien.goede@nwu.ac.za.

methodologies as prescriptive theories. Since the application field is data warehousing, section 4 provides a short introduction to this field of study. Section 5 introduces the method for exploring the descriptive advantages of prescriptive theories. Guidelines for the use of this method of analysis are proposed in section 6.

Descriptive and Prescriptive Theories

As this paper concerns theory, the term will be discussed briefly before more detail is provided on the classification of theories.

What is a theory?

Different disciplines use the concept “theory” to describe different constructs. The concept of “theory” is often used in scientific literature to describe explanations and predictions of phenomena that is testable (Gregor, 2006). Scholars in Information Systems (IS) research have a more social perspective on reality. Lee (1999) describes an interpretative view of the concept. He views theories as results of research and suggests that they are social constructions which are invented rather than discovered.

Midgley (2000) describes the change of how people viewed a theory from a rule governing phenomena to a means of explanation. A theory can be seen as a “way of seeing that explains things in terms of particular purposes and values” (Midgley, 2000). A detailed explanation of positivistic and interpretative views of theory can be found in Gregor (2006). The interpretative view of theory is supported in this paper.

Categorisation of theories

In order to use a specific method or theory in the way it was intended by the authors it is sometimes useful to categorise methods and theories. Gregor (2006) provides a classification for Information Systems (IS) theories in terms of the primary goal of the theory. Gregor identifies the different goals of theories as: Analysis and description; Explanation; Prediction; and Prescription. Table 1 contains a summary of the classification of theory by Gregor (2006) according to the goal of the theory.

Table 1 Theory types in Information Systems (quoted from Gregor (2006))

Theory type	Distinguishing Attributes
Analysis	Says what is. The theory does not extend beyond analysis and description. No causal relationship among phenomena is specified and no predictions are made.
Explanation	Says what it is, how, why, when and where. The theory provides explanations but does not aim to predict with any precision. There are no testable propositions.
Prediction	Says what is and what will be. The theory provides predictions and has testable propositions but does not have well developed justificatory causal explanations.
Explanation and Prediction	Says what is, how, why, when, where, and what will be. Provides predictions and has both testable propositions and causal explanations.
Design and Action	Says how to do something. The theory gives explicit prescriptions (e.g. methods, techniques, principles of form and function) for constructing an artefact.

A similar classification outside the field of IS can be found in Rooke *et al.* (2009). This paper concerns the difference between descriptive theories (the first 4 in table 1) and prescriptive theories (the last category in table 1). Koskela (2000) describes prescriptive theories as guides to doing things in the world. This paper explores the question whether prescriptive theories may be useful when the researcher is not in a position to dictate action. It is done by means of a test case in information systems. The prescriptive theories of systems thinking are under investigation in the field of data warehousing.

Systems thinking methodologies

Systems thinking emerged in reaction to reductionism, when Von Bertalanffy (1968) advocated an interdisciplinary approach to widen the scope when studying problem situations (Ackoff, 1971). A system is a set of interrelated entities, of which no subset is unrelated to any other subset and has properties that do not exist in the parts but are found in the whole (called emergent properties) (Weinberg, 1975).

Many authors, such as Checkland and Poulter (2006), Churchman (1968), and Garajedaghi (2011), stress the purposeful nature of a system. Checkland and Poulter (2006) warn that everything we call in the everyday sense a “system” is not a “system” in terms of systems thinking. They argue that for something to be called a system in the systems thinking use of the concept, there must be some purposeful activity and that systems include communication processes, control processes, structures in layers, and emergent properties (Checkland & Polter, 2006). Churchman (1968) describes systems in terms of their objectives, environment, resources, components, and their management. He argues that a specific system can be identified by its objectives. Gharajedaghi (2011) describes systems in terms of system principles of: openness, emergent properties, purposefulness, multidimensionality, and counter-intuitiveness. His work is related to that of Churchman (1968) since he uses the same definitions for the systems and the environment of the system. Everything that the participating actors in the system can control is part of the system and that which is not under the control of the participating actors is the environment (Gharajedaghi, 2011). Ulrich (1987) highlights the role of a third group, namely the people that do not form part of the participative actors but are affected by the results of the system.

Systems thinking principles are applied to a variety of fields today. Leveson (2012) applies systems thinking ideas to safety. She uses systems theory in terms of hierarchical systems and boundary judgments to propose a more holistic view on situations when safety regulations are designed. Gharajedaghi (2011) advocates the use of systems thinking ideas in designing business architecture.

Different ontological views of systems, which we call methodologies, developed over time namely hard systems, soft systems, and critical systems thinking. This paper follows Jackson's (1991a) broader use of the term methodology whereby methodology refers to methods for exploring and gaining knowledge about systems. It is therefore prescriptive in nature. There are mainly three systems thinking methodologies, namely hard systems thinking, soft systems thinking, and critical systems thinking. Each of these strands of systems thinking is based on a different ontological view of the world. Hard systems thinking is based on realism, soft systems thinking is based on relativism, and critical systems thinking is based on critical social theory.

The development of systems thinking can be seen as an example of Kuhn's description of paradigm development. According to Kuhn (1996) a paradigm is "universally recognized

scientific achievements that, for a time, provide model problems and solutions for a community of researchers." He further writes (1996:12) that "successive transition from one paradigm to another via revolution is the usual developmental pattern of mature science." In terms of systems thinking, scholarly thought moved away for classical operational research where solutions were proposed for specific problems, towards a systemic view world where the world was viewed in terms of wholes (hard systems thinking). Thinking further progressed to a view where worldviews dictated the understanding of a problem situation (soft systems thinking) to current critical thinking that extends soft systems thinking by enlarging the boundary of the system to also include those unrepresentatives affected in the system, sometimes called the oppressed.

Hard systems thinking

Hard systems thinking is a term used by Checkland (1981) as an alternative to "Soft Systems". According to hard systems thinking, social systems are treated like scientific problems. A 'system' becomes a label for something that exists outside us (Checkland, 1999).

From a hard systems perspective a system is viewed as a hierarchically organised set of elements. When one understands the components of the system, one is able to understand the system as a whole. Hard systems thinking focuses on objectivity and ignores issues of subjectivity (Jackson, 1991b). A system is seen as a true representation of reality.

Jackson (1991b) proposes from the work of Checkland (1981) that hard systems thinkers make the following assumptions: 1. There is a desired state of the system, which is known; 2. There is a present state of the system; 3. There are alternative ways to get from the present state to the desired state of the system; 4. It is the role of the systems analyst to determine the best way to get to the desired state of the system. Many other methodologies can be simplified to fit this pattern.

Hard systems thinking in Information Systems Development

From the 1950's Information Systems development was focussed on solving large mathematical problems. Information system development was guided by what was later known as the Waterfall method, described by Benington as the stage wise model in 1956. The phases of this method are: requirements gathering, design, implementation, verification, and maintenance. The aim was to support the Operational Research (OR) practitioner.

Ackoff (1979) describes five reasons for the decline of OR. These also hold for the typical IS systems of that era. The first reason for the decline in OR is the need for learning and adaptation. Computer programs took so long to develop – often they were developed after the OR solution were designed – that by the time they were implemented the organisational setting changed in such a way that they had very limited use.

A second reason for the decline of OR is the omission of aesthetics. Ackoff argues that not enough attention was given to the means or process which produced the ends or result, in terms of individual preferences and styles. This holds especially for information systems where the users were not involved in the process after requirements collection. The development of a suitable user-interface was the responsibility of the technical programmer and users were trained to use it.

A third reason for the decline of OR is interrelatedness of problems. He argues that managers need not solve individual problems but has to deal with "messes" and that the solution of each problem in the mess does not solve the mess. IS systems were also designed

individually for the use of different departments in the organisation. Even today we are battling with data integration from systems where the design did not take into the account other systems in the organisation.

Ackoff argues that traditional OR aims to predict reality and then tries to prepare for it. The focus was on the cause-effect relationship and Ackoff argued that if we can prepare for the future we can affect it. The fourth reason was that OR focused on predicting and preparing for reality with limited focus on shaping reality in future. He argued that the aim of OR should be on designing a “desirable future” and then developing methods to achieve this. Requirements of IS were also affected by this attitude of managers. Computer systems were developed to model current environments. They were not used as tools to present creative new handling of problems.

The final reason for the decline of traditional OR is the disciplinarity of OR. Ackoff argues that by developing OR as a discipline it lost its ability to fully understand the problem situation from different perspectives. Computer systems were also developed from the perspectives of a specific discipline.

One of the major problems with IS in this period was the belief that the requirements for a new computer system could be obtained at the beginning of the project and then be used to design and develop a system that is successful. Success was measured against the requirements and not the actual use of the computer system to aid the original intention.

These elements of critique was taken seriously and many OR researchers today developed OR into a very useful tool in decision making.

Soft Systems thinking

Hard systems thinkers view the world as systemic while soft systems thinkers view the world as filled with complexities and confusion but the process of inquiry is systemic (Checkland, 1981). According to soft systems thinking a system is viewed as a person’s perception of the real world. It follows an interpretative view where the social world is seen as “being the creative construction of human beings” (Jackson, 1991b:126). Different subjective views of the problem environment enhance the understanding of the problem situation. People understand systems from their own point of view or their own worldview which is formed by their underlying assumptions of the world.

Various authors developed methodologies (prescriptive theories) for practicing soft systems thinking. Ackoff (1981) proposed the following phases: 1. Formulating the mess; 2. Ends planning; 3. Means planning; 4. Resource planning, and 5. Design of implementation and control. Churchman (1970) proposes a methodology consisting of phases for 1. Thesis; 2. Antithesis, and 3. Synthesis. The best known methodology for practicing soft systems thinking is that of Checkland (Checkland, 1981; Checkland and Scholes, 1999; Checkland and Poulter, 2006). The Soft Systems Methodology (SSM) is a learning cycle consisting of understanding the real world problematical situation; building purposeful activity models of the situation, which are representative of different world views; comparing the models to the real world situation to aid structured discussion about change; and taking action to improve the situation, which might trigger another cycle (Checkland & Poulter (2006).

Soft systems thinking in Information Systems Development

The information systems development community took note of the SSM. Mathiassen *et al.* (1991) developed a methodology called Rapid Systems Modelling (RSM). When describing

this methodology, they explicitly link it to the SSM and the work of Checkland. They describe how each aspect of the SSM can be translated to the use in IS development. Their focus is on learning about the system and the expected use of the computer system. They achieve this learning by means of prototyping. A prototype can be defined as a “small-scale, representative, or working model of the user’s requirements or as a proposed design for an information system. Any given prototype may omit certain functions or features until such a time as the prototype has sufficiently evolved into an acceptable implementation of requirements” (Whitten *et al.*, 2004:772). This definition has a strong iterative notion which can be linked to the iterative nature of the SSM.

The influence of SSM in IS development was acknowledged when Iivari *et al.* (1994) listed it amongst the five IS development approaches along with structured approach, information modelling, socio-technical approach and object-oriented approach.

Critical systems thinking

Thomas and Locket (1991) argues that the application of soft systems methodologies “increases the power of managers and experts rather than equalising power relationships.” Affected people do not always fully comprehend their position in society due to a lack of knowledge. They cannot represent themselves adequately in a problem solution environment such as SSM and it is the role of the “critical theorist to employ a social theory capable of explaining the alienated world and actions of oppressed groups in society” (Jackson 1991a).

Some critical systems thinkers (such as Jackson, 1991) stress their belief that the world is not fundamentally harmonious. Therefore, to understand, explain and make possible changes, one must think in terms of contradictions. Different perceptions can be seen as expressions of irreconcilable conflict and power struggle between management and workers, or systems developers and users (Dahlbom and Mathiassen, 1993). Intervention is central to practicing critical systems.

There are several methodologies for practicing critical systems thinking. Flood and Jackson’s Total Systems Intervention (TSI) is widely used. It uses metaphors to encourage creative thinking. A discussion can be found in Flood and Jackson (1991). Ulrich (1983) has done extensive work in the field of boundary judgement and provides a method called critical heuristics consisting of 12 questions on affected and involved parties. His work is of such importance that it will be republished in December 2012 in the Encyclopedia of Operations Research and Management Science (see Ulrich, 2012). Midgley (2000) provides details on critical systemic intervention. All these methodologies are more complex than a simple list of phases and needs to be presented in more detail than the scope of this paper. They are however all concerned with identifying and overcoming oppressive structures in problem situations in terms of identifying the plight of unrepresented but affected parties (such as ordinary citizens or future generations).

Critical systems thinking in IS development

A number of authors (Panagiotidis & Edwards (2001); Córdoba & Midgley (2003); Córdoba (2008); and others) have written on the influence of critical systems thinking in IS. Most reports are on specific projects.

One such project was done in IS systems planning by Córdoba and Midgley (2003). They state that IS planning often focuses on aspects such as successful implementation in a given period of time and solving problems by using the latest technology which is troublesome in

especially developing countries. They argue that there are other social aspects to take into consideration. In 2009 Córdoba provided a framework for IS planning which is sensitive to the possibility of the marginalization of some of the involved or affected parties in the problem situation. He does this by using boundary critique of Ulrich (1983) to identify the stakeholders in the problem environment and autopoiesis to guide interaction among stakeholders. In terms of boundary critique they use the notion of Churchman (1968) where the boundary of a problem should include the knowledge and people relevant for analysis of a social design.

Donaires (2006) discusses how the use of the boundary ideas of Ulrich (1983) can be used to gain an improved understanding of information systems development. He gives a discussion of the “is” and “ought to” boundary judgement questions answered for the software development process. This analysis aids understanding of the motivations of software development and usage. It demonstrates the relative lack of control the development team has on the actual usage of the systems. In data warehousing this is especially true. The data warehousing team’s main objective is to supply management with information they need to support their decisions. The development team does not know exactly what kind of queries will be done on the data warehouse and cannot be accountable for decisions made by management, as long as the data in the data warehouse is accurate.

Data warehousing

This paper concerns the application of systems thinking methodologies in data warehousing development. Data warehouses are information systems that are used in parallel with everyday operational systems to provide business decision makers with information on an organisation-wide scale. The importance of data warehousing in the information technology industry is undeniable. Ramamurthy *et al.* (2008) writes: “Data warehousing (DW) has emerged as one of the most powerful technology innovations in recent years to support organization-wide decision making and has become a key component in the information technology (IT) infrastructure”. This is supported by the fact that Gartner (a leading IT research and consulting company) provides chief information officers with explicit advice on trends in data warehousing in 2011 and 2012 (<http://www.gartner.com>). More detail on the nature of data warehousing is provided here to demonstrate the complexity of these systems.

Introduction to Data Warehousing

Inmon (1996) defines a data warehouse as a “subject oriented integrated, non-volatile, and time variant collection of data in support of management decisions.” Kimball *et al.* (1998) simply defines a data warehouse as “the queryable source of data in the enterprise.” Sen and Sinha (2005) performed an extensive investigation into the use of DW development methodologies. They studied 15 different data warehousing methodologies in terms of various aspects. In their discussion of data warehouse tasks and in their description of the aspects of the different methodologies, they refer to concepts developed by either Inmon (e.g. 1996) or Kimball (e.g. Kimball *et al.* 1998). This highlights the importance of the work of these authors. Both of these authors focus on industry literature and their ideas are presented in books rather than peer-reviewed papers. Many interviews with them are published on the World Wide Web.

Inmon and Kimball differ on various concepts in data warehousing, one of which is the development lifecycle of a data warehouse. Inmon (1996) advocates a lifecycle that he calls the CLDS (reverse of SDLC: systems development lifecycle) with the following phases: 1. Implement data warehouse; 2. Integrate data; 3. Test for bias; 4. Program against data; 5. Design DSS system; 6. Analyze results; 7. Understand requirements. This is a data-driven lifecycle methodology. Kimball *et al.* (1998) advocates the use of a requirements-driven lifecycle methodology. His methodology begins with a data warehouse readiness test. Then user requirements are gathered, followed by modelling, data staging, end-user application design, and maintenance. The aim of this section is to give the reader background knowledge on data warehousing without focussing on different strategies. In their comparison of data warehousing methodologies Sen and Sinha (2005) provides a good introduction to the concepts. They also only refer to the work of Inmon and Kimball in this discussion, highlighting the importance of these authors.

The aim of the data warehouse is to give end-users (mostly managers) easy access to data in the organisation. In order to do this, it is necessary to capture everyday operational data from the operational systems of the organisation. Operational systems are transactional systems, for example point of sale systems that are designed around relational databases, which form the source systems of the data warehouse. The data from the source systems go through a process called data staging to the presentation servers (Kimball *et al.*, 1998). Data staging involves four very important actions. Firstly, the data is extracted from the source systems. The data required for the data warehouse is usually distributed in various different source systems with different file formats running on different hardware and operating system platforms. Secondly, the data is transformed to the data warehouse format. Errors in the data and inconsistencies are removed during this phase. Thirdly, the data is loaded into data marts in the presentation server. The final task of the data staging area is to schedule this process. Kimball (2008) replaced the term data staging with the ETL- Process (Extract, Transform, and Load) as this abbreviation became part of industry- jargon.

The presentation server is the heart of the data warehouse. Data marts are stored here (according to Kimball *et al.* (1998); Inmon's implementation of a data mart differs slightly). Data marts are representations of business areas in the organisation. Data is stored as star schemas consisting of fact and dimension tables. Kimball *et al.* (1998) proposes the use of what they call "conformed dimensions". These are data tables that store data concerning entities in the organisation. The design of these tables to satisfy the needs of all business process' users, speeds up delivery of new versions meeting new requirements as the original integration of the data is done with the entire organisation's structure in mind. Kimball *et al.* (1998) used what they call a "bus architecture" matrix to plan the original effort in terms of the scope of the entire organisation. When the data is organised in data marts in the presentation server, it can be accessed with end-user tools.

Access methods differ greatly between operational systems and data warehouses. In operational systems, fixed access methods are pre-built as standardised reports. Users use the data in a predetermined way. In data warehouses, very few standardised reports are written. Users use browsers and *ad hoc* queries to access the data. Data in the data warehouse cannot be altered by the end-users because of the historical nature of the data. However, it is possible to add some of the report outputs of the end-users into data marts to enhance the data warehouse's functionality. These are usually results from data mining that are stored in analytical data marts. Because of the difference in the use of the term "data mart" between

Inmon (1996) and Kimball *et al.* (1998), Kimball *et al.* (2008) refrained from using the term as he argues that the term has been marginalised by others narrowing its meaning.

Different authors identify success factors in data warehouse design. Inmon (1996) states: “Building data marts before developing a data warehouse can be one of your biggest mistakes”. Mimno (2001) argues that the most important success factor is to make your data warehouse business-driven. He argues that a technology-driven approach is much more likely to fail than a business-driven approach.

In a study on success factors Sammon and Finnegan (2000) identify what they call ten commandments of DW success:

1. A business-driven data warehousing initiative
2. Executive sponsorship and commitment
3. Funding commitment (budgeted and unexpected) based on realistically managed expectations
4. Project team with access to cross-functional project management and implementation experience
5. Attention to source data quality
6. A flexible enterprise data model
7. Data stewardship
8. A long term plan for automated data extraction methods / tools
9. Knowledge of DW compatibility with existing systems
10. Hardware/software proof of concept

These are very much supporting the ideas of DW readiness proposed by Kimball *et al.* in 1998. Other studies on data warehousing success such as Shin (2008) and Hwang & Xu (2008) re-enforce these ideas.

Motivation for using systems thinking in data warehousing

Data warehouses are systems that integrate various data sources to supply management information in organisations. Many data warehouse (DW) projects take longer than originally planned and cost much more than initially budgeted for. The Cutter Consortium reported that 41% of practitioners surveyed have experienced data warehouse failures (Anon., 2003, www.cutter.com). However, the benefits of successful data warehouses are so significant that the academic research community should search for methods to improve the success rate of data warehouse projects.

The author of this paper was drawn to systems thinking as a basis for improving data warehouse quality because of the holistic nature of systems thinking. Kevin Strange (2001) of Gartner wrote: “With respect to the analytical (BI) side of customer relationship management, at least 65 per cent of the efforts are implemented in an unintegrated fashion, based on a function (different efforts by different departments), rather than on a more strategic initiative – the sum is larger than the parts.” Many authors (Eckerson, 2003; Mimno, 2001; Hwang & Xu, 2008) argue that business objectives should be central to DW planning and development. This is very much in line with the systems approach proposed by Churchman (1968). He advocates that subsystems should work together to achieve the objectives of the system, and the objectives of the subsystem should relate to that of the system. A question worth investigating is whether systems thinking can provide more pointers to successful data warehouse development practices. Wixom and Watson (2001)

argue that data warehousing success factors can be grouped in three categories: organisational, project and technical success factors. Whang and Xu (2008) demonstrate how operational, technical, economical, and scheduling factors combine to impact on the benefits for the organisation. From a systems thinking perspective this is an indication of the boundary of a data warehouse as well as an indication that one needs collaboration from different sections in the organisation, which will lead to different worldviews.

Checkland, who described soft systems thinking as a reaction to hard systems thinking, was influenced by the work of Vickers (1965). Vickers moved away from the idea of goal seeking models towards the idea of appreciation as value judgement. Checkland (1985) describes Vickers' idea of an appreciative system as a "cultural mechanism which maintains desired relationships and eludes undesired ones." Vickers advocated "relationship maintaining" as the basis of a richer, more realistic model for organisational decision making than goal seeking (Checkland, 1985). A similar idea is proposed by Ashurst *et al.* (2008) who advocate that benefits of an IS system to the organisation should be more important than technical success:

Perhaps most importantly organisations need to move away from considering the successful delivery of a new piece of software as being the primary objective of a systems development project, and concentrate on the delivery of real business benefits, which might only be realized once users begin to appropriate the technology and adapt it to their own requirements and working contexts.

Systems thinking methodologies are prescriptive in nature. They are designed to guide actions. As data warehousing is a very expensive activity, very few data warehousing managers would be prepared to attempt a new method that they do not have knowledge of. The author of the paper wants to establish to what extent data warehouse professionals' thinking relates to specific systems thinking methodologies intuitively, without having knowledge of these methodologies. This information would enable the researcher to make the implicit relationship between systems thinking and data warehousing explicit. By doing this kind of analysis it can be demonstrated to managers that their team's current practices are to some degree similar to a specific systems thinking methodology. Such a mapping will enhance their understanding of the intrinsic motivations of their team members, which in turn might lead to them agreeing to adopt a specific methodology prescriptively in conjunction with their technical methods or at least understand differences in the thinking of members of their teams.

Using systems thinking methodologies to describe data warehousing practices

This section introduces a mapping method for understanding a phenomenon – in this instance data warehousing – from the perspective of a prescriptive theory – in this instance systems thinking methodology. As indicated before it is not realistic to expect a large organisation to develop a data warehouse according to a methodology unknown to them. A research project was done to establish whether the actions of data warehousing practitioners can be linked to systems thinking methodologies.

Three case studies were conducted in different industries. The first was done in the health industry. The team investigated is responsible for the integration of a hospital linked pharmacy network's data into a data warehouse. The second team investigated is part of the information management of a large bank, aiming to provide management with strategic information based on banking transactional activity of clients. The third team consisted of

management and technical staff of a data warehousing consulting firm who provides data warehousing solutions to a wide range of listed organisations.

People on different levels in the organisational hierarchy were interviewed. In all the cases studied, senior managers were interviewed. Front room data warehousing staff responsible for business analysis requirements gathering were interviewed. In all the case studies back room data warehousing staff, responsible for technical implementation, were interviewed.

Semi-structured interviews were conducted to gather data. Questions were posed to the data warehouse development teams on six major data warehousing aspects: data warehouse adoption, data warehouse development methodology, requirements collection, data modelling, data staging (including data quality) and end-user applications.

Nineteen interviews in total involving 12 staff members (4 managers, 2 project managers, 3 front room staff members, and 3 back room staff members) were held, each lasting between one and three hours. Interviews were recorded and transcribed.

The aim of the case studies was to determine whether the thinking leading to practices of data warehouse professionals can be linked to specific systems thinking methodologies. Prior to analysing the data a preliminary mapping between systems thinking methodology and data warehousing practice was developed.

A specific misconception of the method may be to use the method to “find” systems thinking activities in the practices of the data warehouse practitioners and then to declare the actions as representative of systems thinking. Such an approach is flawed as the method is designed to find systems thinking activities – it will be found. The aim is rather to discover in which way people think about systems, not that they are thinking about systems. The reader is reminded of the evolution of thought about systems as a process from hard to soft to critical systems thinking.

Mapping between systems thinking and data warehousing

Creating a mapping between systems thinking and data warehousing is a difficult but creative process as no literature on the topic could be found at the time. The mapping was developed iteratively, by first describing how a specific systems thinking methodology will be used prescriptively to develop a data warehouse. This was done from data warehouse literature, combined with systems thinking literature. When the descriptions of the different methodologies applied to data warehousing were compared differences were highlighted. These differences became the aspects that were questioned in the interviews.

The complete mapping consists of 105 questions on data warehousing with possible responses from each systems thinking methodology. Table 2 is an extraction of the mapping used to analyse the case data. The complete table can be found in Goede (2004). For each question, a typical answer was formulated in terms of each of the systems thinking methodologies investigated. No other connection between systems thinking methodologies and data warehouse practices could be found in literature, and the author hopes to make a contribution to the systems thinking community through this mapping.

Questions were phrased to be open-ended and the options stated in Table 2 were not provided to the respondents. The answers in table 2 were used by the interviewer as guidelines during the interviews. Table 2 indicates how seemingly similar practices are diversely motivated and ontologically rooted. This table can also be used to determine the methodological viewpoint in an organisation.

Case study representation and analysis

The aim of the case studies was to indicate that different data warehousing practitioners' motivations are rooted in different systems thinking methodologies, although their practices may seem similar. Table 3 contains an excerpt of how the mapping of Table 2 was used to analyse the case data. It is not intended to depict the complete data set.

Each answer given by each respondent was mapped onto a specific cell in the table by comparing the answer given by the participant with the answers in Table 2. Each respondent was identified by an abbreviation of his/her function within the organisation. In this example IM was the Information Manger, SP the Staging Programmer, DM the Data Warehouse Manager, and OM the Operations Manager.

Three tables were created representing each of the organisations studied in the respective case studies. Two of the most important problems were to map answers that do not correspond directly to Table 2 and to decide how many follow-up questions to ask. Footnotes were used to provide the exact answer or a summary thereof to address the first problem. The decision of how many follow-up questions to ask remained a problem and was handled for each individual by the researcher while trying not to lead the respondent.

The analysis of the data was done in several iterations to ensure accuracy. A time gap of about 30 days was enforced between the analysis iteration. Ideally more than one person should be responsible for the data analysis, this however was not possible. Data analysis was done in the following sequence:

1. The mapping table was completed about 60 days prior to the first interview.
2. Each interview was transcribed and checked. During the evaluation period the entire interview was listened to in order to obtain a better understanding of the silences and emotions before each question's answer were analysed.
3. In a table similar to table 2 the complete answer of a specific question of a specific participant was copied. The answer was compared with the precompiled answers in each column. Aspects of the answer that could be linked to the precompiled answers were identified and the identifying code of the participant was added to the corresponding cell in a master table (similar to table 3) for each case study. In the example table given here (table 3) the following identification codes were used: WM (warehouse manager), IM (information manager), DM (data manager) and SP (staging programmer). If the answer did not correspond directly but the meaning of the answer indicated a specific way of thinking a foot note was added that contained the reasoning behind the allocation.
4. Step three was very time-consuming. After all the answers were analysed, some minor changes were made to the mapping to be more representative of typical answers, which led to a repetition of the entire analysis process. Since this research was conducted on a part-time base and considering the exhaustive nature of the analysis, about 30 days passed between analysis iterations.
5. At the start of each iteration blank tables were used and the process was repeated from the beginning without looking at the previous analysis.
6. After an iteration of analysis was completed, the differences between iterations were investigated, often leading to the updating of footnotes.
7. Only when a next iteration did not involve any changes were the analysis deemed as completed.

The hermeneutic nature of interpretive data analysis was taken into account. The analysis was conducted answer by answer per participant after the entire interview was listened to.

The idea of changing the mapping table during analyses, sparking off further iterations of analysis – is representative of the development of a method while using that method. Checkland and Howell (1998) provide a diagram that explains this well (Figure 1).

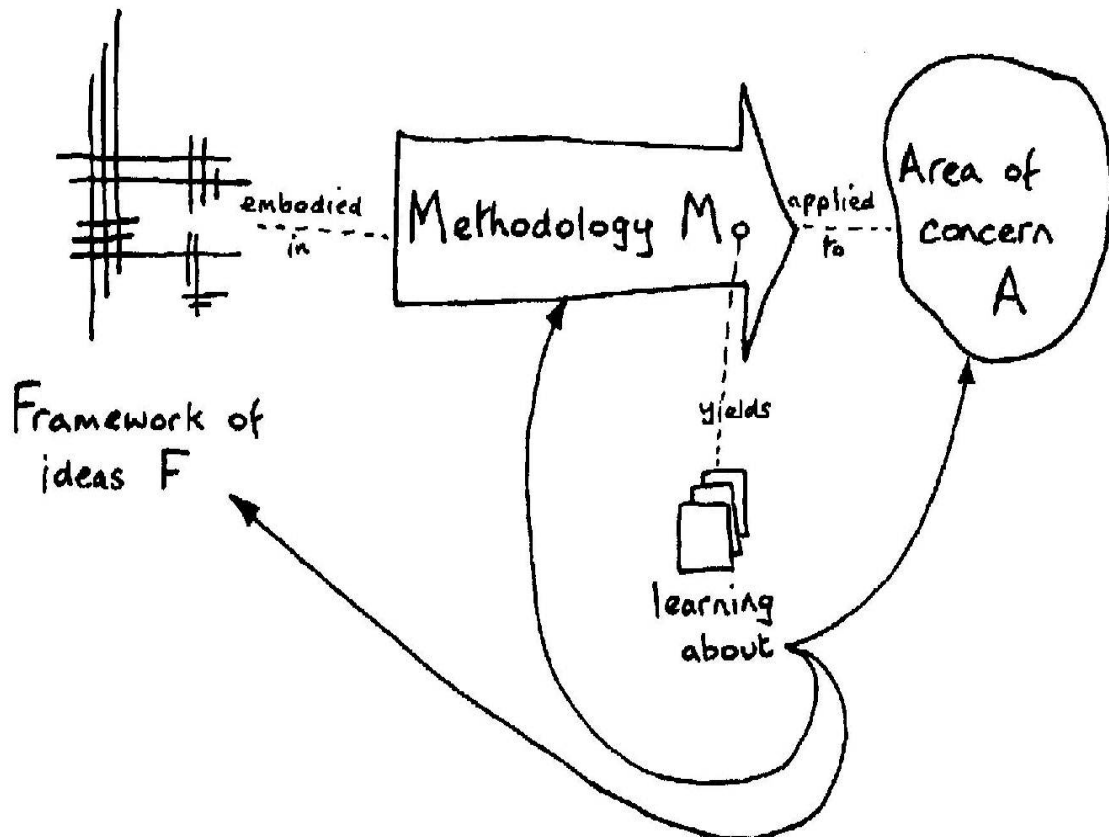


Figure 1 Elements relevant to any piece of research (Checkland and Holwell, 1998)

They use this diagram to demonstrate the development of SSM. In this study the diagram can be used to describe the development of this method of mapping ideas. The framework of ideas (F) is the evolution from hard to critical systems thinking applied to data warehousing. The area of concern (A) is the everyday actions of data warehousing practitioners represented by the interview content. The data analysis using the mapping table, can be seen as the

methodology (M). While using the methodology learning occurs on all three aspects represented by F, M and A.

Data analysis yielded a table per case study (similar to table 3) indicating to which extent the thinking of the participants could be understood from the perspective of a specific systems thinking methodology.

Learning from the results of the analysis

The analysis tables (similar to table 3) were constructed row by row in terms of the answers of specific individuals of specific questions. The completed tables should be interpreted from a column perspective. If a specific column has many entries it implies that when one listens to the team member's answers from the perspective of systems thinking methodologies, specific systems thinking methodology's ideas were frequently voiced by the participants, without them being aware of these methodologies. From table 3 provided as an example it is clear that not many answers could be perceived as critical systems thinking answers in the portion of data represented by this table.

The completed tables in this study yielded different results in terms of thinking about data warehousing from a systems thinking perspective. The first organisation (in the banking sector) from which the analysis provided in table 3 is an excerpt had very few answers mapped in the critical systems thinking column. The project leader and the manager responses could mostly be linked with soft systems thinking. Their technical programmer often gave answers linked with hard systems thinking. In the discussions of the results with the manager, he was interested in the difference of answers of specific questions – which represented contentious issues within the team. This highlighted the advantage of this kind of analysis. This manager expressed that he better understands a specific team member's work by looking to the table entries for this team member from a column perspective. He also highlighted entries where his code and that of the project leader were in different cells.

The answers of the participants of the second organisation (in the health sector) were more linked to the soft systems thinking than was the case in the first organisation. There were more responses linked to critical systems thinking than in the first case, but the majority of responses were allocated to soft systems thinking. The interest of the manager in understanding the completed tables was similar to that of the first case's manager.

Very few answers of the third organisation's participants were linked to hard systems thinking. A larger number of their responses to the asked questions could be linked to critical systems thinking than in the second case study. As this organisation is in the consulting industry their focus is strongly on measurable benefits to specific sections in larger organisations, which might explain the higher connection to critical systems thinking. The managers had a similar interest in the completed analysis tables as the managers from the other cases. As two managers were interviewed they found the comparison of the allocation of their answers very interesting.

Table 2 Example from detailed mapping between systems thinking methodologies and data warehousing

	Question to data warehousing team member	Hard systems thinking	Soft systems thinking	Critical systems thinking
SECTION A: Data warehouse adoption				
1	What is a data warehouse?	A data warehouse is an integrated data source to fulfill the reporting needs of business units. It consists mainly of data, metadata, and technology such as computers.	A data warehouse is a system to improve decision making in the organisation. It consists of people, data and technology.	A data warehouse is a tool to affect positive change in the organisation as a whole. It consists of everything required to succeed in the realisation of the proposed change.
2	What is the root problem to be solved by the data warehouse?	Data quality issues when providing information to management – same version of the truth.	To aid the organisation's strategic objectives.	To solve a specific problem in the organisation through active intervention or change.
3	Who owns the data warehouse?	The development team.	More than one party, but mostly the users.	Both the involved and the affected.
4	What is the impact of the data warehouse on other systems or business?	Not sure, mostly technical.	Impact study was performed. Overall data quality is improved.	Groups that were previously regarded as outside the data warehouse are now part of the data warehouse depending on the scope or boundaries.
SECTION B: Data warehouse development methods				
1	Describe the lifecycle of the development of the data warehouse.	Inmon's lifecycle or a strong waterfall model where user participation is limited to the requirements collection phase.	Strong focus on user participation and the organisation's objectives. Definitely an iterative process.	This would be an iterative approach, but more than one methodology could be used according to the applicability. There is a critical awareness of the weaknesses of each methodology.

	Question to data warehousing team member	Hard systems thinking	Soft systems thinking	Critical systems thinking
2	What is the role of end-users in the development lifecycle of a data warehouse?	End-users are responsible for the requirements definition. They are not involved in the development process. They are involved in the final stages of testing of the system.	End-users are involved in as many phases as possible, including requirements collection and modelling, as well as end-user application development. An incremental process is used to accommodate end-user views in the system.	End-users are motivated by their specific goal and involve themselves to ensure their goal is achieved. Participation and emancipation go hand in hand. End-users are on equal status in the data warehousing team. They are central to the boundary judgment of the project.
SECTION C: Requirements definition				
2	Do users know what they want? How do you go about assisting them?	No, users don't know, but we deliver typically what they ask us to do. It is their problem. Alternatively, an Inmon approach, where requirements are developed later in the project, after the technical implementation has been completed.	Not always, but the data warehousing team should help them to specify their different views, all of which are combined into user specifications on the basis of consensus.	Mostly yes. The group initiating the data warehouse has strong motivation for the development of the data warehouse. It is the task of the data warehousing team to be critical towards them in order to identify power struggles and negative intentions toward another group or individual in the organisation.
SECTION E: Data staging and data quality				
1	Who is responsible for data quality?	The source system owners as far as availability and accuracy are concerned and the IS team as far as compatibility is concerned.	A joint effort. The source systems should be in line with the organisation's objectives and should therefore be willing to adapt to achieve objectives. The users should indicate which data definitions are most representative of the organisation's practice.	High quality source data is part of a successful data warehouse. The data warehouse team ought to have a representative of the source system. Source systems that provide poor quality data need to be changed.
SECTION F: End-user applications				
1	How do you know when the data warehouse is successful?	When data quality improves and when the specifications are met.	When the organisation's objectives are better achieved.	When the intervention is achieved.

Table 3 Representation of data from one case study

	Question to data warehousing team member	Hard systems thinking	Soft systems thinking	Critical systems thinking
SECTION A: Data warehouse adoption				
1	What is a data warehouse?	DM gave a technical description. WM IM ¹ SP	OM ²	
2	What is the root problem to be solved by the data warehouse?	SP	DM: Management Information for high level decision making. WM: Better results for the organisation. IM	
3	Who owns the data warehouse?	SP WM	IM: "The whole organisation." IM ³ OM gave the name of the organisation as answer	DM: "Business."
4	What is the impact of the data warehouse on other systems or business?	IM ⁴ SP	OM ⁵	
<p>*The abbreviations IM, SP, WM, OM refer to team members and are discussed in section 5.2</p> <ol style="list-style-type: none"> 1. "In short I would say it is a centralised location for data on different roll-ups....." 2. "In theory, people should take the organisation's objectives into account when building a data warehouse, but in practice very few of them do." 3. The data marts belong to business, but the technical data store belongs to the information management department. 4. The impact is very limited, perhaps here and there a quality issue. 5. The source system's quality should improve according to the work done by the data warehousing team; there should be a closed feedback loop. 				

Generalization and guidelines of using of this method

After this method was used to better understand the systems thinking orientation of data warehouse practitioners, the generalisability of it were conceptually evaluated. The first step is to identify specific characteristics of the situation where it was developed. The data warehousing professionals did not have any knowledge of systems thinking. One could not deductively achieve understanding of their actions from a systems thinking perspective by

using typical deductive research methods such as grounded theory. The analysis of the answers would produce a theory on how data warehouses were developed from a data warehousing perspective and not from a systems thinking perspective.

Another characteristic is that systems thinking methodologies are linked to specified methods or methodologies which these practitioners did not follow explicitly. In general these methods can be viewed as prescriptive in nature.

After identifying the characteristics which motivated the development of the method, other situations with similar characteristics could be identified.

One such situation arose in the personal environment of the author. The author studied the work of Maria Montessori on raising children. Although this work inspired many aspects that are central to preschool child care today very few pre-school teachers are aware of this work. The methods proposed by Montessori are prescriptive in nature. Pre-school teachers do not appreciate parents advising new methods! The method proposed in this paper can be used to set up a table where two columns are used, one for a typical “Montessori” answer and the other to indicate an absence of “Montessori” related answer. The table can then be used to indicate which current teaching strategies reflect Montessori teaching. Change which is an extension of current behaviour is more likely to be implemented.

Once again, it is important to note from a research perspective that such research is not aimed to prove that pre-school teachers use the ideas of Montessori, but rather to show that when their practices are viewed from a Montessori perspective some similarities can be found.

This method can also be used to study the cognitive models used by pupils when learning a specific skill such as computer programming. A mapping will be created of how different cognitive models will approach certain aspects of problem solving required in programming.

This method can be used in all cases where the behaviour of participants is viewed from a point of view of which the detail is unknown to the participants.

The author of this paper believes that this method is of great use in studies with an interdisciplinary character since the researcher uses the mapping table to combine aspects of two different disciplines not previously combined.

The following guideline should be followed when this method is used:

1. The mapping table should only be created after a clear literature review on both the areas of interest. It is useful to have an additional column to the mapping table with the motivation from literature for asking a specific question.
2. When setting up the mapping table for a research project where methodology is used to understand practice, the underlying philosophical assumptions of the methodology should be used to guide the application of the methodology in the application area. In this case the ontological assumptions of systems thinking methodologies were used to apply systems thinking methodologies to data warehousing practices.
3. The table must be verified rigorously before data collection is done and the participants should not have access to the content of the table. This would ensure more intuitive answers. If the participants have the table they might want all their answers to “fit” in one column.
4. Data analysis should be done iteratively – to increase the rigor of the process, it is advisable to do the analysis thoroughly and then after a week or so to do it again

without looking at the first attempt. When comparing the results care should be taken to go back to the data to re-analyse the differences.

5. It is helpful to arrange follow up sessions with the participants after the analysis – this however should be done with care as not to ask leading questions to the participants. Remember not all answers have to be mapped in the table.
6. From an interpretative research point of view one should always state your own preferences prior to the analysis of the data. Klein and Meyers (1999) provide more detail on this and other aspects of interpretative research in IS. In this case the author believes that a hard systems approach to all aspects of data warehousing may have a higher possibility of data warehouse project failure, but certain technical issues are accommodating of hard systems thinking.

Summary and conclusion

Prescriptive theories dictate action in specific situations. One is not always able to prescribe how an action is to be performed. This might be due the cost of an experiment or the lack of knowledge of the role players in the situation. This paper explores the possibility to use prescriptive theories descriptively. This is done by creating a mapping of how certain prescriptive theories would address a specific situation. Actual data is then fitted onto this mapping in order to determine to what extent the prescriptive theory fits the current state of events. This mapping can then be used to guide future action by highlighting either differences or similarities with the current state of events and the prescribed theory.

Future research can be done to refine the guidelines provided in section 6 to improve research projects where prescriptive theories are used to describe actions of practitioners.

References

- Ackoff, R.L. (1971). Towards a system of systems concepts. *Management Science*, 17(11): 661-671.
- Ackoff, R.L. 1981. *Creating the corporate future*. New York, N.Y.: Wiley.
- Anonomous, (2003). "41% have experienced data warehouse project failures", Cutter consortium report of 18 February 2003. [Web:] <http://www.cutter.com/research/2003/edge030218.html>. [Date of use: 13 June 2012]
- Ashurst, C., Doherty, N.F. and Peppard, J., (2008). Improving the impact of IT development projects: the benefits realization capability model. *European Journal of Information Systems*, 17 (4): 352-370.
- Checkland, P. (1981). *Systems thinking, systems practice*. Chichester: Wiley.
- Checkland, P. (1991). From optimizing to learning: A development of systems thinking for the 1990 (*In Flood, R.L. & Jackson, M.C., eds. Critical systems thinking: Directed readings*. Chichester: Wiley: 59-75.).

- Checkland, P. & Holwell, S. (1998). Information, systems and information systems. Chichester: Wiley.
- Checkland, P. & Polter, J. (2006) Learning for action. Chichester: Wiley.
- Checkland, P. and Scholes, J. (1999). Soft systems methodology in action: Includes a 30-year retrospective. Chichester: Wiley
- Churchman, C.W. (1968). The systems approach. New York, N.Y.: Delta.
- Churchman, C.W. (1970). Operations research as a profession. *Management Science* (17): B38-B53.
- Dahlbom, B. & Mathiassen, L. (1993). Computers in context: The philosophy and practice of systems design. Oxford: NCC Blackwell.
- Donaires, O.S. (2006). A critical heuristic approach to the establishment of a software development process. *Systemic Practice and Action Research*, (19): 415–428.
- Eckerson, W. (2003). Smart companies in the 21st century: The secrets of creating successful BI solutions. (In A report of the data warehouse institute) [Web:] <http://download.101com.com/pub/TDWI/Files/V8I3TDWINewsletter.pdf> [Date of use: 13 June 2012].
- Ferranti, M. (1998). Checking out success, failure, and myth. *Data warehouse world*, Aug. [Web:] <http://sunsite.uakom.sk/sunworldonline/swol-08-1998/swol-08-datawarehouse.html>. [Date of use: 13 June 2012]
- Flood, R.L. & Jackson, M.C. (1991) Total systems intervention: A practical face to critical systems thinking. (In Flood, R.L. & Jackson, M.C., eds. Critical systems thinking: Directed readings. Chichester: Wiley: 19-40.)
- Gharajedaghi, J. (2011) Systems thinking: Managing chaos and complexity: A platform for designing business architecture (3rd ed). Burlington, M.A.: Morgan Kaufman.
- Goede, R. (2004). A framework for the explicit use of specific systems thinking methodologies in data-driven decision support system development. Ph.D. dissertation. University of Pretoria, Pretoria. Available at: <http://upetd.up.ac.za/thesis/available/etd-05132005-080727/>
- Gregor, S. (2006). The nature of theory in information systems, *MIS Quarterly*, 30(3): 611-642.
- Hwang, M.I. & Xu, H. (2008). A structural model of data warehousing success, *The Journal of Computer Information Systems*, 49(1): 48-56.
- Iivari, J., Hirscheim, R. & Klein, H.K. (1999). Beyond methodologies: Keeping up with information systems development approaches through dynamic classification. *Proceedings of the 32nd Hawaii international conference on system sciences*: 1-10.
- Inmon, W.H. (1996). Building the data warehouse. 2nd ed. New York, N.Y.: Wiley.
- Jackson, M.C. (1991a). Systems methodology for the management sciences. New York, N.Y.: Plenum.

- Jackson, M.C. (1991b). Social systems theory and practice: The need for a critical approach (In Flood, R.L. & Jackson, M.C., eds. *Critical systems thinking: Directed readings*. Chichester: Wiley: 117-138.)
- Kimball, R., Reeves, L., Ross, M. & Thornthwaite, W. (1998). *The data warehouse lifecycle toolkit*. New York, N.Y.: Wiley.
- Kimball, R., Ross, M., Ross, M., Thornthwaite, W., Mundy, J. & Becker, R. (2008). *The data warehouse lifecycle toolkit*. 2nd ed. New York, N.Y.: Wiley.
- Klein, H.K. & Myers, M.D. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS Quarterly*, 23(1): 67-94.
- Koskela, L. (2000). An exploration towards a production theory and its application to construction, VTT Technical Research Centre of Finland, VTT Publication 408.
- Kuhn, T. S. (1996). *The structure of scientific revolutions*, 3rd ed. Chicago, Ill.: Univ. of Chicago Press.
- Lee, A.S. (1999). Researching MIS. (In Currie, W.L. & Galliers, R., eds. *Rethinking management information systems*. Oxford: Oxford University Press: 7-27.)
- Leveson, N. (2012). *Engineering a safer world: Systems thinking applied to safety*. MIT, Mass: MIT press..
- Midgley, G. (2000). *Systemic intervention: Philosophy, methodology and practice (contemporary systems thinking)*. New York, N.Y.: Kluwer Academic/Plenum.
- Mimno, P. (2001). Is your data warehouse business-driven? Published at [Web:] <http://www.mimno.com/articles/article3.html>, [Date of use: 13 June 2012].
- Ramamurthy, K., Sen, A. & Sinha, A.P. (2008). Data warehousing infusion and organizational effectiveness. *IEEE transactions on systems, man, and cybernetics—part a: systems and humans*, 38 (4): 976-994.
- Rook, J.A., Koskela, L., & Kagioglou, M. (2009). Informality in organisation and research: a review and a proposal. *Construction Management and Economics* (27): 913-922.
- Sammon D. & Finnegan, P. (2000). The ten commandments of data warehousing. *The data base for advances in information systems*, 31 (4).
- Sen, A. & Sinha, A. P. (2005). A comparison of data warehousing methodologies. *Communications of the ACM*, 48(3): 79-84.
- Shin, B. (2008). An exploratory investigation of system success factors in data warehousing. *Journal of the association for information systems*, 4: 141-170.
- Strange, K. (2001). Data warehouse: Essential infrastructure for BI. [Web:] <http://www.gartner.com/id=348959> [Date of use: 13 June 2012].
- Thomas A.R. & Lockett, M. (1991). Marxism and systems research: Values in practical action (In Flood, R.L. & Jackson, M.C., eds. *Critical systems thinking: Directed readings*. Chichester: Wiley: 85-102.)
- Ulrich, W. (1983). *Critical heuristics of social planning: A new approach to practical philosophy*. Bern: Haupt.

- Ulrich, W. (2012). Critical systems thinking (*In* Gass, S. I. & Fu, M.C eds. Encyclopaedia of operations research and management science, 3rd Edition, London: Springer, publication pages not yet available.
- Vickers, G. (1965). The art of judgment. London: Chapman & Hall.
- Von Bertalanffy, L. (1968). General system theory: Foundations, development, applications. (Revised *ed.*) New York, N.Y.: George Braziller.
- Weinberg, G.M. (1975). An introduction to general systems thinking. New York, N.Y.: Wiley.
- Wixom, B.H. & Watson, H.J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly*, 25(1): 17-41.