THE INNER AND INTER CONSTRUCT ASSOCIATIONS OF THE QUALITY OF DATA WAREHOUSE CUSTOMER RELATIONSHIP DATA FOR PROBLEM ENACTMENT

A thesis submitted in partial fulfilment for the Degree of Doctor of Business Administration

by

Raul Mario Abril

Henley Management College
Brunel University

December 2005

THE INNER AND INTER CONSTRUCT ASSOCIATIONS OF THE QUALITY OF DATA WAREHOUSE CUSTOMER RELATIONSHIP DATA FOR PROBLEM ENACTMENT©

I certify that all material in this dissertation which is not my own work has been identified and that no material is included for which a degree has previously been coferred upon me.

Raul Mario Abril Jimenez

Abstract

The literature identifies perceptions of data quality as a key factor influencing a wide range of attitudes and behaviors related to data in organizational settings (e.g. decision confidence). In particular, there is an overwhelming consensus that effective customer relationship management, CRM, depends on the quality of customer data. Data warehouses, if properly implemented, enable data integration which is a key attribute of data quality. The literature highlights the relevance of formulating problem statements because this will determine the course of action. CRM managers formulate problem statements through a cognitive process known as enactment.

The literature on data quality is very fragmented. It posits that this construct is of a high order nature (it is dimensional), it is contextual and situational, and it is closely linked to a utilitarian value. This study addresses all these disperse views of the nature of data quality from a holistic perspective. Social cognitive theory, SCT, is the backbone for studying data quality in terms of information search behavior and enhancements in formulating problem statements.

The main objective of this study is to explore the nature of a data warehouse's customer relationship data quality in situations where there is a need for understanding a customer relationship problem. The research question is *What are the inner and inter construct associations of the quality of data warehouse customer relationship data for problem enactment*?

To reach this objective, a positivistic approach was adopted complemented with qualitative interventions along the research process. Observations were gathered with a survey. Scales were adjusted using a construct-based approach. Research findings confirm that data quality is a high order construct with a contextual dimension and a situational dimension. Problem sense making enhancements is a dependent variable of data quality in a confirmed positive association between both constructs. Problem sense making enhancements is also a high order construct with a mastering experience dimension and a self-efficacy dimension. Behavioral patterns for information search mode (scanning mode orientation vs. focus mode orientation) and for information search heuristic (template heuristic orientation vs. trial-and-error heuristic orientation) have been identified. Focus is the predominant information search mode orientation and template is the predominant information search heuristic orientation. Overall, the research findings support the associations advocated by SCT. The self-efficacy dimension in problem sense making enhancements is a discriminant for information search mode orientation (focus mode orientation vs. scanning mode orientation). The contextual dimension in data quality (i.e. data task utility) is a discriminant for information search heuristic (template heuristic orientation vs. trial-and-error heuristic orientation).

A data quality cognitive metamodel and a data quality for problem enactment model are suggested for research in the areas of data quality, information search behavior, and cognitive enhancements.

Page iv missing in original

Acknowledgements

I cannot express in words my deep feelings of gratitude and emotion at this stage. I will fail to acknowledge someone or something. I am probably completing one of the most important steps in my life and, looking retrospectively, it could not have happened without the help of a number of people and institutions who have supported me in this research. I am deeply grateful to my employer Teradata, a division of NCR, for the partial financial support to this study. The many professional opportunities in which I have been involved have given me a unique perspective on the subjects of this research. My supervisors, Professors Bob O'Keefe and Joe F. Hair, Jr. and Dr. David Price are the best team that I could have dreamt of. In general, thank you all for the privilege of being my supervisors, for your feedback and guidance. In particular, Bob, I appreciate your 'keep-it-simple' recommendations and knowledge on information systems research, thanks Bob. Joe, I appreciate your expertise in the analysis and your help in 'squeezing' the data from so many perspectives, not mentioning your contribution to the questionnaire design, thanks Bob. David, thanks a lot for your guidance since my first day in the DBA program, and for your patience with my demanding requests, thanks David.

I am grateful to a number of scholars who have helped me: Professor Pat Joynt, thanks a lot for your incisive comments on methodological aspects and for teaching me not to say "I think that ...", thanks Pat. Professor Frank Pajares, thanks a lot for your input in the self-efficacy construct and scale, gracias Frank. Professor Wagner A. Kamakura, thanks for your feedback in relation to CRM and the questionnaire, thanks Wagner. Professor Teresa Brannick, thanks a lot for your input on the scales, thanks Teresa. Professor Kay Wolf, thanks for your input in relation to the adoption of you self-efficacy scale, thanks Kay.

I must recognize the mountain of past research illustrated by the 2,968 entries in my bibliographic database. Thanks to all the authors for allowing me to contribute, as Dr. David Price says, with a grain of sand. Also, I have to recognize the help of the successive librarians of my employer Mary Hill, Heidi Lock, Cecilie Rask, and

Fatimah Ahmed. Mary, Heidi, Cecilie, and Fatimah with your skills and time I collected 1,629 prints and 728 e-files, thanks a lot for this massive source of literature.

Likewise, I appreciate the insightful comments I received from fine practitioners like Ole Darre Christrup, Francisco Siso, Jake Johnson, Frank Davis, Sam Sterling, Dr. Graham Doig, Dr. Les Bowd, Dr. Mats Lindgren, Dr. Ralf Müller, and Dr. Per Olaf Brett, thanks a lot. A big thanks to all the people around the world who participated in the focus groups and provided input to the questionnaire.

Thanks to SAS for the free of charge software licenses of Base SAS, SAS/STAT, SAS/INSIGHT, and SAS Enterprise Miner.

Thanks a lot to my college, Henley Management College, for funding the use of the TeleForm software for my survey and for the great support of Gill Baker and Karen Whiting in the implementation of the web-questionnaire and administration of downloads. Thanks a lot to the DM Review publication for supporting this study in their web edition. Also, thanks to the following seven associations for their support: DataWarehouse.com (in their web-portal), The Data Management Association (in their web-portal and in their monthly communication to members), The Project Management Institute Information Systems Special Interest Group (e-mail to members), Colegi Oficial d'Engineers en Informatica de Catalunya (e-mail to members), Canadian Information Processing Society (monthly communication to associates), Asociacion de Tecnicos en Informatica (e-mail to members), and The Global Billing Association (monthly communication to members).

And my greatest thanks are saved for my mother who never wavered in her support. I dedicate this dissertation with love to her. She deserves a doctorate honoris causa, one earned for wisdom, devotion, energy and love. Gracias mama por habermelo dado todo ...y mas. ¡Juntos lo hemos conseguido!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 MOTIVATION FOR THE STUDY

I wanted to learn and explain the nature of data quality in situations where it is necessary to understand customer relationship problems. Consequently, one must focus on the individual's cognition and behavior rather than the technological aspects of data. My motivation results not just from the outcome of these research findings but also from the knowledge acquired in the research process and the expectation of being able to bring to bear other professional and academic endeavors.

My education in computer science and my professional experience implementing decision support systems based on information technology since 1985 made me feel reasonably comfortable with the literature that deals with this type of subjects. In one way or in another, the literature that addresses information systems and their data repositories, for example data warehouses, in organizational settings highlights the instrumental nature of such systems in a given context.

Some research (e.g. [Wierenga and Van Bruggen 1998]) contends that the success of implementations of information systems will depend on the fit between the system and personal factors of their users. In simple terms, it is suggested that information systems should be designed considering the individual's traits rather than expecting users to adapt to the support systems (e. g. [Davis, Bagozzi and Warshaw 1989]).

Social and psychology sciences place the individual on both sides of the equation and not just on one side as, in general, it is the best case scenario in engineering sciences with the notable exception of the recent usability subject in design and testing (e.g. [Butler 1996]). Donald E. Knuth [2001], probably the guru of gurus in computer science, contends that sciences will evolve to a super specialization characterized by disciplines that cover two bodies of knowledge.

Hence, intellectual growth will require communication based on a network of dyadic associations [Knuth 2001]. In this book, Professor Knuth presents the common

ground shared by computer science and theology. From my anecdotal experience, it is quite a challenge to move from computer science to literature that deals with cognition and human behavior. For example, it is fascinating to realize that the same concept construed in information theory (computer science) as entropy [Shannon and Weaver 1949] and in cognitive fit theory [Vessey 1991] (psychology science) as data quality require such different learning skills in order to be able to grasp their meaning.

Therefore, it was of relevance for this research to have (i) an integrative approach to data quality considering key aspects like cognitive aspects and behavior and (ii) to deconstruct data quality in order to gain conceptual clarity.

1.2 DATA, INFORMATION, KNOWLEDGE AND BELIEFS

The term data refers to symbols obtained through an encoding process of the environment (e.g. [Burke 1989]) that can be available to an individual but which have not as yet been evaluated for their worth in a specific situation-within-context (e.g. [McDonough 1963]). In business settings data is stored in data files (e.g. data warehouses) and retrieved by individuals engaged in information search behaviors. I have represented these concepts in Figure 1.1.

The retrieved data is interpreted in specific situations and contexts. In the literature on cognition (e.g. [Burke 1989, Goia 1986, Feldman and March 1981]), information is defined as data with an imparted meaning by an individual through a cognitive process called enactment. Data requires meaning to derive information, and the meaning must stem from a specific [problem] situation-within-context. No matter how much data is processed, it cannot be turned into information until a manager uses it in a given situation-within-context, for example, to solve a problem.

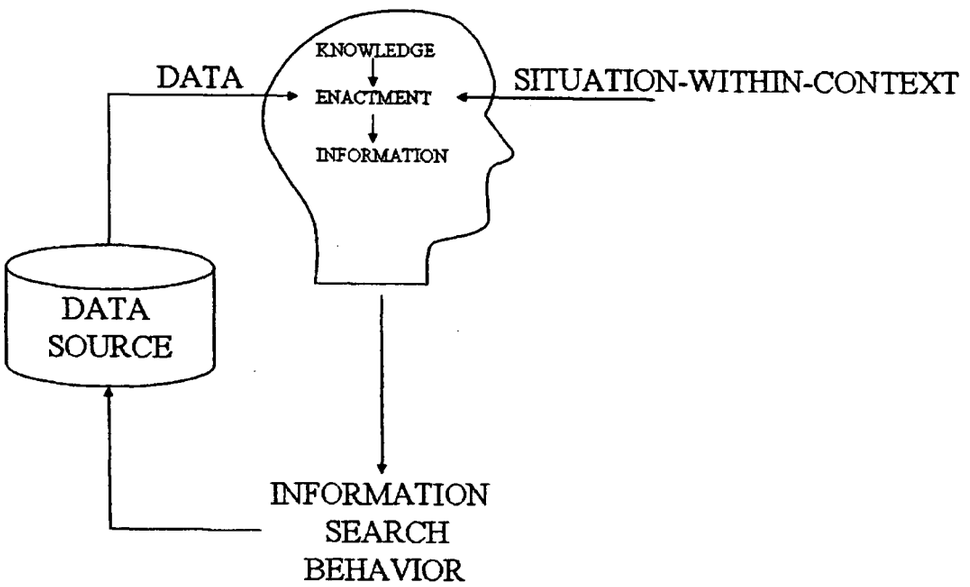


Figure 1.1. Data versus information versus knowledge

Knowledge is information believed by an individual as justified truth (e.g. [Nonaka 1994]) and stored in memory (i.e. it can be retrieved) in a cognitive structure (e.g. [Lamberts and Shanks 1997]) through a cognitive process called learning. By the term cognitive structure or cognition it is meant the mental organized representation of information (e.g. [Cyert and March 1963, March and Simon 1958, Simon 1955, Walsh 1995, Argyris and Schön 1978]). Cognitive structure refers to the manner in which an individual's knowledge is organized (e.g. levels of abstraction, decomposition, causal and functional relationships) while knowledge refers to what information is available. Therefore, according to Wang and Chan [1995] the two concepts complement each other.

For the purposes of this research we adhere to Nonaka's view [1994] in that we focus on the individual's belief about the justification of knowledge and not on its truthfulness (i.e. an individual's knowledge might be falsifiable and not scientifically generalizable). Knowledge refers to what information is available in memory (e.g. [Wang and Chan 1995]). From a historical perspective, there are aspects of theories that "survive" to the next step in the evolution. For example, in the early twentieth century the dominant epistemological approach was logical positivism. Many aspects of this philosophy are still present in cognitive science research. According to this approach, knowledge consists of schemas that attempt to represent the environment in

such a way as to maximally simplify problem-solving. A more radical point of departure is offered by constructivism, which assumes that all knowledge is built up from scratch by the subject of knowledge. The idea of a correspondence or reflection of external reality is rejected. The danger with constructivism is that it may lead to relativism; to the idea that any model constructed by a subject is as good as any other and that there is no way to distinguish adequate or 'true' knowledge from inadequate or 'false' knowledge. We can distinguish two approaches that try to avoid such an 'absolute relativism'. The first may be called individual constructivism. It assumes that an individual attempts to reach coherence among the different pieces of knowledge. Constructions that are inconsistent with the bulk of other knowledge that the individual has will tend to be rejected. Constructions that succeed in integrating previously incoherent pieces of knowledge will be maintained. The second approach, called social constructivism, sees consensus between different subjects as the ultimate criterion to judge knowledge. 'Truth' or 'reality' will be accorded only to those constructions on which the majority of a social group agrees.

A practical aspect of knowledge is its instrumental utility. One must have knowledge of the task demands because activities differ in difficulty (e.g. [Bandura 1997]). Again, situation-within-context plays a determinant role because the same activity taps differing knowledge under different circumstances (e.g. [Bandura 1997]). Human information-processing theories of cognition (e.g. [Anderson 1983, Newell and Simon 1972]) provide a framework to understand the acquisition and transfer of knowledge. Adaptive Control of Thought-Rational (ACT-R) theory of skill acquisition [Anderson 1993] is one of those approaches. ACT-R is based on the assumption that there are two long-term stores of knowledge: declarative knowledge and procedural knowledge.

Declarative knowledge includes facts, instructions, examples and concepts. It is knowledge that we can consciously recall [Anderson 1993]. Procedural knowledge is based on skills to perform a specific task [Anderson 1993]. The term skill refers to a learned cognitive capability and an associated proficiency at performing a task (e.g. [Kanfer and Ackerman 1989]). The term capability refers to the capacity for doing something [Oxford 1993]. According to [Oxford 1993] the terms ability, capability

and competence refer to the same concept. People tend to make attributions of cognitive capability (i.e. self-beliefs) either as an acquirable skill or as a stable entity (i.e. fixed inherent capacity) [Dweck and Elliot 1983].

By the term belief is meant a personal ontological posture about some element of reality (e.g. [Rowland 1995]) formulated as a predicate on a subject cognitively constructed by individuals (e.g. [Pajares 2002]). Beliefs and knowledge are different in several aspects. For example, a belief is concerned with the existence or nonexistence of a certain conceptual entity. Beliefs rely heavily on evaluative and affective components and can also be held with varying degrees of certitude [Abelson 1979].

It is axiomatic in the literature (e.g. decision sciences, information management, marketing) that eliciting information is one of the key activities in the process that enables the enactment of information. The contextual and situational nature of information has been theoretically (e.g. [Weick 1995]) and empirically (e.g. [Brannick 2000]) observed in academic research. Both aspects are of paramount importance because adherence to them has a conflictive implication in accepting research of a general nature on human information processing. For example, research findings on information search behavior that are of a general nature will be severally criticized as unrealistic because of a deficiency in the definitions of the construct studied due to the missing aspects of contextual and situational nature.

1.3 PROBLEM ENACTING

The term problem refers to the individual's perception of a variance, or a gap, between the present and some desired state of affairs (e.g. [Simon 1977, Smith 1990]). Structured problems are structured because we choose to treat them as such. Such critique is important because it highlights the relevance of defining a problem. For the purpose of this paper, the term problem statement refers to the formulation in linguistic terms of the problem elements and its structure (e.g. [Smith 1989, Abualsamh, Carlin and McDaniel Jr. 1990, Pitz, Heerboth and Sachs 1980]). Problem statements are cognitively construed. Cognitive processes refer to the mental

processes involved in the acquisition, organization and use of information (e.g. [Bandura 1994]).

Enactment is a genuine contribution of the sense making theory (e.g. [Weick 2000]). Sense making is a high order cognitive process intended to reduce equivocality, or multiple meanings, in the information (e.g. [Weick 2000]). For example, making sense of a customer relationship problem means that heedful interrelating connects sufficient individual knowledge with situational demands [Weick 1993]. By enactment is meant the generation of information, plausible interpretations of a (problematic) situation, and actions to be realized (e.g. [Weick 2000]).

Because sense making is grounded in identity construction and based on plausibility, we might expect differences owing to differences in the individual [Taylor 1999]. Because sense making is enactive of sensible environments and focused on extracted cues, we might expect individuals in different environments to make sense of things differently [Taylor 1999]. Identity construction, plausibility, and a particular set of extracted cues also seem tied to the characteristics of a particular job and we might expect people with different jobs to make sense of data differently [Taylor 1999].

1.4 OBJECTIVES OF THE STUDY

There are two basic definitions that are of key importance in this research: The definition of data warehouse and the definition of customer relationship management. A data warehouse stores data from the task environment [Witteloostuijn 1996]. A data warehouse is a repository of integrated data that comes from the many transactional systems that support the mission-critical business processes in the organization and from external data sources (e.g. [Marakas 1998]). Customer relationship management, CRM, is the cross-functional business process that drives customer value by the creation and maintenance of business–to-customer durable, close and mutually beneficial relationships (e.g. [Leigh and Marshall 2001]).

The main objective of this study is to explore the reflective nature of the quality of data warehouse customer relationship data in situations as a latent variable where

there is a need for understanding a customer relationship problem. This main objective involves a number of secondary objectives, for example, the consideration of data quality as both the independent and dependent variable with respect to cognitive and behavioral variables. Before this study, I expected that data quality would be associated to cognitive aspects, to be identified in my literature review, and to information search behaviors, likewise to be identified in my literature review. Also, the only expected association before my literature review was that data quality and cognitive improvements are associated. This expectation was based on anecdotal experience and "popular wisdom in the industry".

One aspect of understanding problems, and therefore part of this research, it is the formulation of a problem statement. However, outside the scope of this research is any solution aspect of the problem. In simple terms, this study focuses on the understanding of customer relationship problems using customer relationship data extracted from a data warehouse as opposed to solving such problems.

The study of the associations between data quality and both cognitive and behavioral variables is also part of this research. Furthermore, the impact of the CRM job function and the number of supported CRM data warehouse functions moderating the associations and/or creating group differences is also part of this research.

1.5 RESEARCH FOCUS

The terms context and situation refer respectively to the set of tasks and situations under consideration in research studies. The types of situations in Figure 1.2 are adopted from the literature in decision/sense making (e.g. [Abril 2001]).

Making sense of customer relationship problems involves a number of constructs and processes. What are the indicators of constructs like data quality and sense making enhancements? Do they interrelate with each other? How? Are they strong? Are there information search patterns? The reviewed literature suggests an affirmative answer; however, most of it neglects information search behavior.

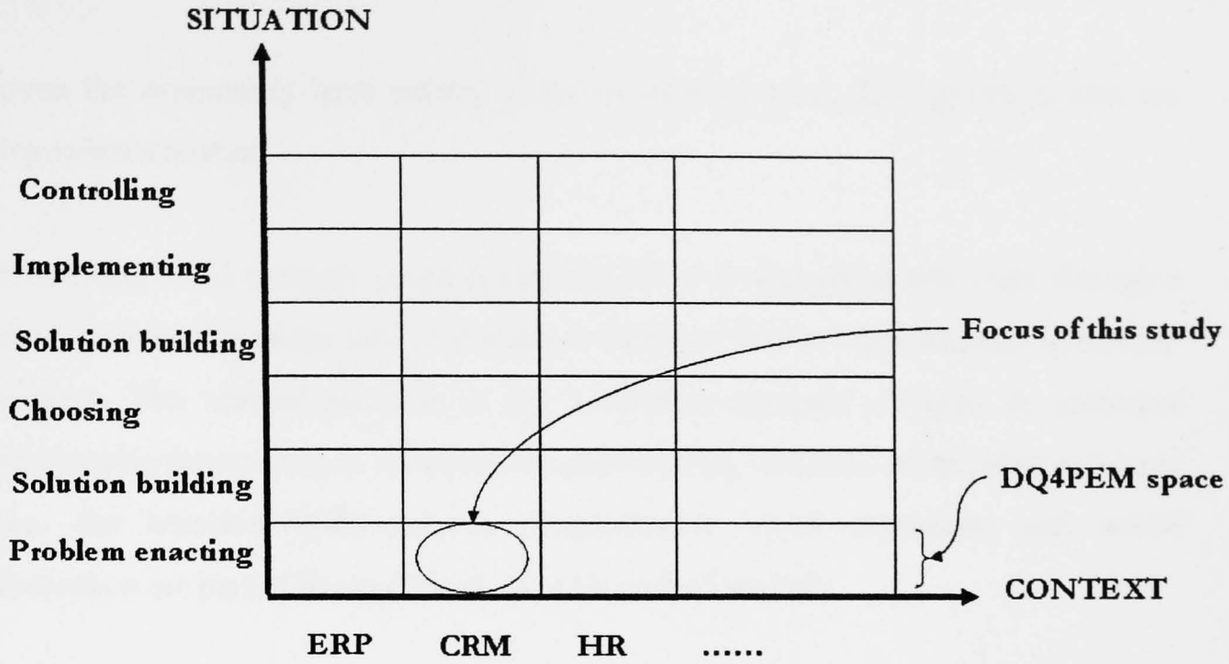


Figure 1.2. Focus of this study in terms of a business context and a business situation

Research in business settings is also typically contextualized around business processes (e.g. manufacturing, human resources, marketing). CRM is the context where this study takes place. Is it true that quality enhances problem enactment, as theorized in the literature, in a CRM context? We are interested in finding responses to these questions in a context of CRM supported by a data warehouse and in a situation that focuses on customer relationship problem enactment.

Our research model (see Figure 3.1) focuses in terms of a context and a sitituation as it is clearly indicated in Figure 1.2. As a result of our research results (see Figure 1.3) the DQ4PEM model is suggested for further research (see Figure 7.4).

Therefore, we have formulated the following research problem statement: The inner and inter construct associations of the quality of data warehouse customer relationship data for problem enactment is not well understood. Our research question is: *What are the inner and inter construct associations of the quality of data warehouse customer relationship data for problem enactment*? This research question belongs to the management information systems evaluation category of questions, which is concerned with the dependent variable in marketing information processing research [DeLone and McLean 1992]. This means that we are considering data quality as the cornerstone of this research (i.e. the independent variable). In fact,

given the exploratory implications of my research question, data quality is also the dependent variable.

In addition to the research question, research focus is achieved in this study through a clear specification of the unit of analysis, a business context and a situation within the context. The unit of analysis is the individual manager engaged in customer relationship processes (i.e. the context) understanding customer relationship problems (i.e. the situation-within-context). Requirements about experience and actual dedication are part of the specification of the unit of analysis.

1.6 RESEARCH APPROACH AND RESEARCH RESULTS

A positivisticic approach was adopted. This approach was complemented with qualitative interventions along the research process. Scales were adjusted using a construct-based approach using focus groups and factor analyses. A pilot intervention was performed before launching the large scale survey. Observations were analyzed using quantitative techniques. Data analyses required six consecutive phases. Association between data quality and sense making enhancements were confirmed using regression anaylisis and structural equation modeling. Research findings were validated in semi-structured interviews.

Results of this study (see Figure 1.3) confirm that data quality is a high order construct with a contextual dimension and a situational dimension. Problem sense making enhancements is a dependent variable of data quality in a confirmed positive association between both constructs. Problem sense making enhancements is also a high order construct with a mastering experience dimension and a self-efficacy dimension. Behavioral patterns for information search mode (scanning mode orientation vs. focus mode orientation) and for information search heuristic (template heuristic orientation vs. trial-and-error heuristic orientation) have been identified. Focus is the predominant information search mode orientation and template is the predominant information search heuristic orientation. The self-efficacy dimension in problem sense making enhancements is a discriminant for information search mode orientation (focus mode orientation vs. scanning mode orientation). The contextual dimension in data quality (i.e. CRM data task utility) is a discriminant for information

search heuristic (template heuristic orientation vs. trial-and-error heuristic orientation). Overall, the research findings support the associations advocated by SCT (e.g. [Bandura 1997]).

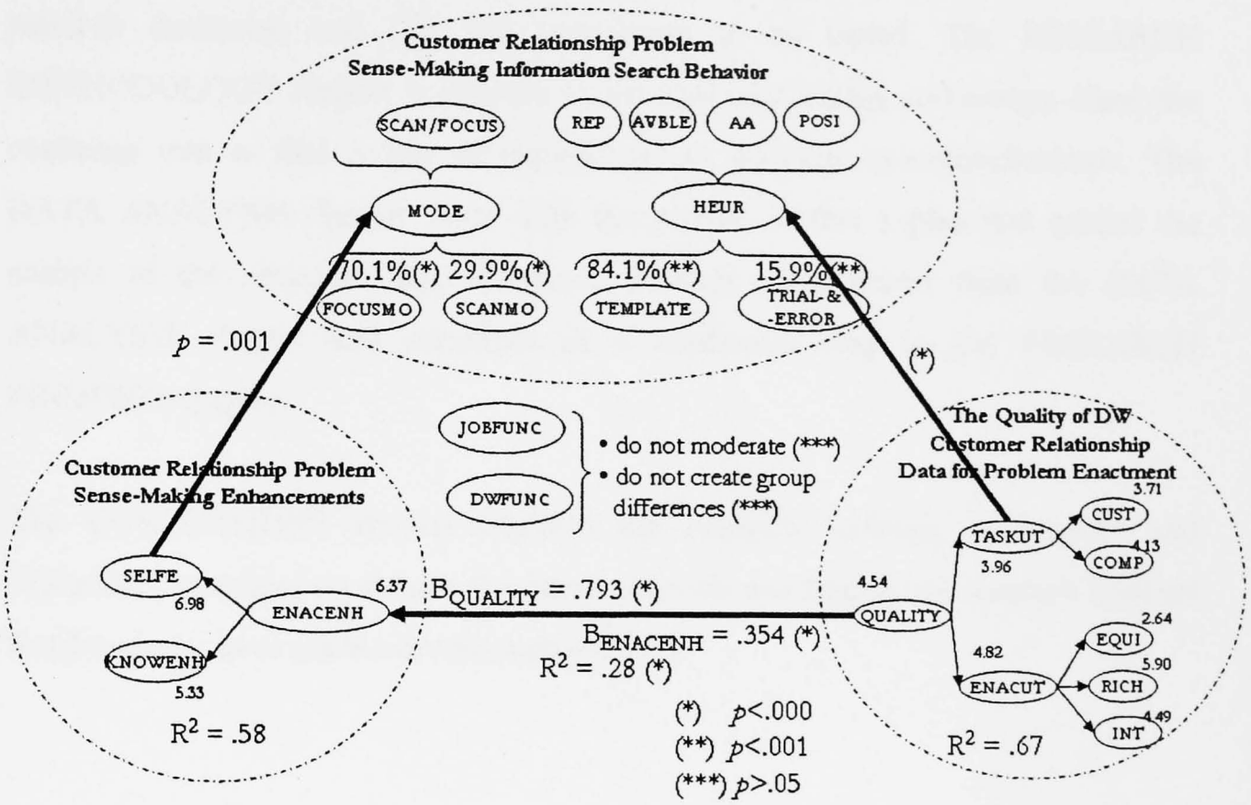


Figure 1.3 Research results

As a conclusion of this study, a data quality cognitive metamodel is suggested for research in the areas of data quality, information search behavior and cognitive enhancements. In addition, as an instance of this metamodel, a data quality for problem enactment model is suggested for research in problem enactment situations. This metamodel and this model are contextually generic and support the triadic associations posited by SCT (e.g. [Bandura 1997]) in situations of problem enactment.

1.7 ORGANISATION OF THE STUDY

The content of this thesis follows the traditional approach for positivisticic research. The LITERATURE REVIEW chapter was designed to facilitate its utilization as a reference material source during the reading of the rest of this thesis. That is, the sections in this chapter map the constructs and links of the research model. The research problem statement and the research question are in this chapter. The last appendix, appendix O, has all the main definitions. The CONCEPTUALIZATION

chapter starts with a description of the context, situation and unit of analysis. After this first section, there are three sections describing (i) the research model including definitions of the higher latent constructs, (ii) the type of research and high-level research decisions, and (iii) the hypotheses to be tested. The RESEARCH METHODOLOGY chapter is devoted to the research strategy and design. Here, the challenge was to find a way of explaining the different operationalizations. The DATA ANALYSIS chapter starts with the phases, in fact a plan that guided the analyis of the observed data. Research findings were drawn from the DATA ANALYSIS chapter and presented in a condensed way in the RESEARCH FINDINGS chapter.

The CONCLUSIONS chapter explains the research findings, weaknesses and limitations, proposes an agenda for future research and frames the research findings that I consider have made relevant contributions.

2. LITERATURE REVIEW

2.1 INTRODUCTION

The objective of this literature review was to find clues for understanding the impact of data quality in a customer relationship management context with a focus in situations where the formulation of a problem statement is required. The term situation-within-context highlights the concept that a situation happens in a context. The layout of this chapter was designed in order to facilitate its utilization as reference material during the reading of the rest of this thesis. That is, the sections in this chapter map the constructs and links of the research model. This literature review follows the paradigm of the substantive foundation of explanations [Sorge 1996], which means that theories can be distinguished according to the constructs they refer to and to the stream of research they adhere to (e.g. utilitarian, socio-cognitive, information processing).

In essence, the steps that I followed in my literature review consisted in a convergent-divergent continuum around data quality until I "discovered" SCT in one of my multiple divergent reviews of cognitive implications. In fact, in my first paper [Abril 2001], where I documented my findings after reviewing the data quality literature, this theory was missing. SCT provided the "building blocks" of my literature review around the constructs that I should consider in order to understand data quality from an holistic perspective.

A challenge in assembling so many conceptual pieces from various theoretical bodies of knowledge is that there are situations where constructs lack conceptual clarity, induce to equivocality (more than one interpretation) and/or are simply not applicable across theoretical bodies. Therefore, I considered it valuable to keep a rigorous record with all the relevant definitions (see APPENDIX O: DEFINITIONS).

The PRACTICAL RELEVANCE OF DATA QUALITY COPING IN BUSINESS SETTINGS section addresses the relevance of data quality from a business perspective. THE RESEARCH SITUATION-WITHIN-CONTEXT: PROBLEM

SENSE MAKING IN CRM USING A DATA WAREHOUSE AS THE SOURCE OF CUSTOMER RELATIONSHIP DATA section introduces SCT and the key constructs in the research situation (i.e. making sense of problems) within the research context (i.e. customer relationship management supported by a data warehouse). This section includes the formulations of the research problem and the research question. Focus in this study is addressed by a research context and a research situation. The research context of this study is described in the RESEARCH CONTEXT: CUSTOMER RELATIONSHIP MANAGEMENT SUPPORTED BY A DATA WAREHOUSE section and the research situation is described in the RESEARCH SITUATION: MAKING SENSE OF PROBLEMS section. Finally, the specific literature about data quality is presented in the CUSTOMER RELATIONSHIP DATA QUALITY FOR PROBLEM ENACTMENT section.

2.2 PRACTICAL RELEVANCE OF DATA QUALITY COPING IN BUSINESS SETTINGS

By data quality coping I mean that data quality can be framed using the same kind of analysis as statistical hypothesis testing, which identifies two types of errors, or low data quality, for the purpose of this research. Drawing on the concepts of types of uncertainty (e.g. [Brannick 1998]), perceived low data quality in a data source can manifest itself in two ways. The first, which I am going to refer as type A data quality, is ignorance on the quality level that is needed for the situation-within-context. The second, which I am going to refer as type B data quality, is being aware of the quality level that is needed for the situation-within-context but being unable to achieve that level. Just as a clarification, the ideal situation (i.e. high data quality) would be to have a very low probability associated with both types of errors.

Without formally entering on the construct definition of data quality, which will be addressed in the RESEARCH DESIGN section, data quality is a popular subject, typically addressed by market analysts in their reports with a focus on type B data quality. In general, these reports anecdotally indicate manifestations of data quality issues and argue about their consequences. Likewise, overwhelmingly, formal academic research takes as a premise that individuals know the data quality that is

needed (i.e. low probability in type A data quality) and focus on type B data quality. This study follows this stream of research.

An intuitive reason for arguing for the relevance of data quality comes from the expectation of the task's demands. In this sense, regarding the marketing effectiveness of a company or division, Kotler [1988] states that marketing effectiveness is reflected in the degree to which it exhibits five major attributes of a marketing orientation: customer philosophy, integrated marketing organization, adequate marketing information, strategic orientation, and operational efficiency. According to this author, each of these attributes can be measured. Therefore, remembering the definitions of data and information in the previous section, data quality seems to play a relevant role in order to have adequate marketing information.

Several market analysts (e.g. [Agosta 2002, Eckerson 2001]) provide an indication of the practical relevance of data quality in their reports. Leaving aside my ignorance about the methodological rigor of such studies, some results are of interest from an anecdotal point of view and, at the same time, raise some questions. For example, the question concerning the toughest challenges facing data warehousing implementations, data quality was listed in first place (18% of respondents) and customer data integration challenges was in sixth place in [Agosta 2002]. This tells us little about the data quality itself; it just says that it is not given. Eckerson [2001] reported that almost half of the companies that were contacted believed the quality of their data was excellent or good, yet almost half of the respondents assessed that the quality of their data was worse than everyone thinks [Eckerson 2001]. What does "worse than everyone thinks" mean when we match it to "excellent data quality" or "good data quality"? It leaves us in a kind of limbo. Nevertheless, in this survey [Eckerson 2001], some reported that specific effects which attributed to low data quality, framed as problems, were extra time to reconciliate data, extra costs (e.g. mail addresses), customer dissatisfaction, and loss of credibility in a system.

Also in [Eckerson 2001], some reported that specific effects which attributed to high data quality, now framed as benefits, were less time spent reconciling data, reduced costs, increased customer satisfaction, and greater confidence in analytical systems.

Unfortunately, looking at the percentages one realizes that while the responses about the effects of low data quality are in the range 67% to 87%, the responses about the effects of high data quality are in the range 12% to 19%. This is a pretty common pattern in this type of report. Scores for the issues attributed to low data quality are higher than the scores for the benefits attributed to high data quality. This suggests that it is easier to find informants about issues on data quality than benefits about data quality.

Are companies reacting to data quality? The overall conclusion in [Agosta 2003] is yes, reporting that only 11% were doing nothing or were in a type A data quality (i.e. ignorance on the data quality level that is needed). Furthermore, firms are spending money in this action, as indicated by a compound annual growth rate of 66 percent in 2001 and 2002 [Agosta 2002]. The list of potential data quality issues is long. Just for explanatory purposes, some examples are:

- Format issues: Probably the most famous data quality issue in history, the Y2K Bug.
- Equivocal issues: A paradigmatic case is the definition of a customer in retail banking. Is a person who is not an account holder and uses an ATM for withdrawing cash considered a customer? There are retail banks that consider a customer to be an account holder and it is not infrequent to find retail bankers that only can estimate the number of customers they have.
- Integration issues: Different records, probably in different data sources, about the same entity (e.g. a customer) with complementary (e.g. fix line calls, cellular calls) or contradictory data (e.g. different addresses). Why has data quality become a big issue in managing customers? In many companies, customer data were originally collected and managed departmentally [Foss, Henderson, Johnson, Murray, and Stone 2002]. This meant bringing data from various systems together, adding new and relevant customer data. These data are typically sourced from several operational or product systems, as well as from marketing, sales and service systems. So interdepartmental systems evolved, with interfaces between them [Foss *et al.* 2002].
- Incomplete issues: Missing values (e.g. blanks) and missing data (e.g. unit of currency).

- Garbage-in, garbage-out issues: Names that are simply incorrect (e.g. "Superman").

2.3 THE RESEARCH SITUATION-WITHIN-CONTEXT: PROBLEM SENSE MAKING IN CRM USING A DATA WAREHOUSE AS THE SOURCE OF CUSTOMER RELATIONSHIP DATA

2.3.1 The Research Problem Statement and the Research Question

From my literature review, it is clear that making sense of customer relationship problems involves a number of constructs and processes. Do environmental constructs like data quality, information search behavior constructs like mode and heuristics and cognitive constructs like knowledge and self-efficacy have an internal structure? If yes, this would add conceptual clarity to their definitions. Do they interrelate with each other? How? In which direction? Are they strong? The reviewed literatures suggest an affirmative answer for all these questions; however, most of them only address dyadic associations, that is between two constructs. Is there a comprehensive model explaining all those constructs and associations? Is it true that quality enhances problem enactment as theorized in the literature? Are there moderators? I am interested in finding responses to these questions in the context of CRM supported by a data warehouse and in a situation that focuses on customer relationship problem enactment. Therefore, I have formulated the following research problem statement: The inner and inter construct associations of the quality of data warehouse customer relationship data for problem enactment is not well understood. My research question is,

What are the inner and inter construct associations of the quality of data warehouse customer relationship data for problem enactment?

This research question belongs to the management information systems evaluation category of questions, which is concerned with the dependent variable in marketing information processing research [DeLone and McLean 1992]. This means that I am considering data quality as the cornerstone of this research (i.e. the independent variable). In fact, given the exploratory implications of my research question, data quality is also the dependent variable.

2.3.2 Social Cognitive Theory

Social cognitive theory [Bandura 1986], SCT in short, provides a comprehensive theoretical underpinning to my research question. The comprehensiveness attribution is relevant because all of the other theories, with the exception of sense making theory, miss at least one of the types of constructs in my research question. Comparing SCT research with sense making research, it is clear that SCT provides a more robust inter-constructs structure than sense making. So, I decided to adopt SCT as the overall theory guiding this research, leaving a supportive role to sense making and all the other reviewed theories.

SCT [Bandura 1986] emphasizes the role of self-efficacy and knowledge as key determinants in behaviour. He supports an agentic view of personality (i.e. individuals are proactive and self-regulating). Individuals' beliefs exercise a self-regulation role by controlling their thoughts, feelings and actions. In the view of SCT, human behaviour is the result of the interplay between these beliefs and the environment. SCT is a bridge between behaviourism and cognitivism. SCT argues that human functioning is the result of a determinism of cognitive, behavioural, and environmental factors and that self-efficacy becomes instrumental to the goals that individuals pursue and to the control individuals are able to exercise over their environments (see Figure 2.1). It does not mean that all the bidirectional associations occur simultaneously nor do they have the same strength [Wood and Bandura 1989].

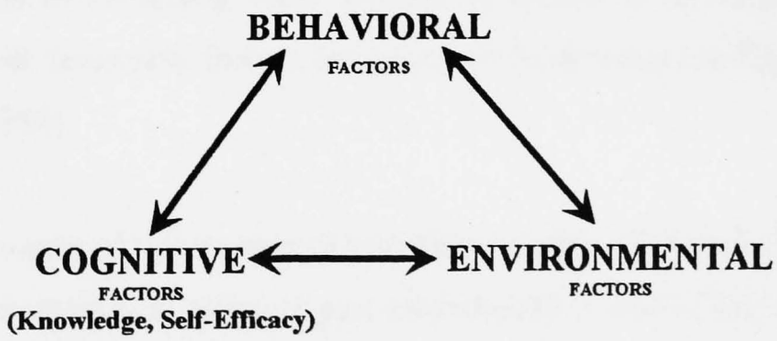


Figure 2.1 Social cognitive theory

Self-regulation of motivation and performance attainments is governed by several self-regulatory mechanisms operating in concert. They include affective self-evaluation and perceived self-efficacy for goal attainment [Bandura 1986]. Perceived

managerial self-efficacy can influence personal goal setting and use of analytical strategies. The stronger the perceived self-efficacy, the higher the goals people set for themselves [Wood and Bandura 1989]. What people know, or the attainments they have previously accomplished in a given environment, are often predictors of subsequent attainments because the beliefs that they hold about their abilities and the outcome of their actions powerfully regulate their behaviour [Pajares 2002] and their effort expenditure accordingly [Klein 1989]. This explains the link {cognitive factors, environment} → behaviour in Figure 2.1. Consequently, how people behave can often be better predicted by their beliefs about their capabilities than by what they are actually capable of accomplishing [Pajares 2002]. This does not mean that people can accomplish tasks beyond their capabilities simply by believing that they can, for competent functioning requires harmony between self-beliefs and possessed skills and knowledge [Pajares 2002]. Stajkovic [1998] did a meta-analysis of 114 self-efficacy published studies finding significant correlation between self-efficacy and task performance.

People are producers of their environment provided that (i) the level of personal efficacy needed to effect changes through enlistment of effort and creative use of capabilities and resources is reached and (ii) how changeable the environment is [Wood and Bandura 1989]. In general, the environment constitutes a potentiality that is actualized by appropriate action. Which parts of the environment are controllable will depend on how people behave. People who have a firm belief in their efficacy figure out ways of exercising some measure of control in an environment which explains the link {cognitive factors, behavior} → Environment in Figure 2.1 [Wood and Bandura 1989].

What are the sources of information of self-efficacy beliefs? Bandura [1986] contends that enactive mastery (i.e. personal past attainments) is one of the most important sources of information and that it is the result of people's behavior in an environment. This explains the link {behavior, environment} → cognitive factors in Figure 2.1. In my literature review on SCT, I found some research mixing SCT with schema theory (e.g. [Brewer and Treyens 1981, Stein 1992]). According to an e-mail

communication from Professor Pajares, probably one of the best experts in SCT, I should just consider the SCT constructs and not schema theory [Pajares 2002].

2.3.3 Cognitive Factors as the Dependent Variables

The link {data quality, information search behavior} → {self-efficacy, knowledge} in Figure 2.1 is supported under the theoretical basis of Bandura's SCT [Bandura 1986], Weick's sense making theory (e.g. [Weick 2000]) and Simon's human information processing theory (e.g. [Simon 1976]). These theories suggest that data quality and information search behavior influence problem enactment both in knowledge (i.e. past attainments) and self-efficacy (i.e. prospective attainments) because:

- Enactive mastery (i.e. personal past attainments) results from people's behavior in an environment (e.g. [Bandura 1986]) and is the most influential source of efficacy information.
- Information acquisition behavior and conveying mechanisms of an organization are key determinants for problem enactment (e.g. [Huber and Daft 1987, Dutton, Fahey and Narayanan 1983]).

The link data quality → {self-efficacy, knowledge} in Figure 2.1 is supported under the theoretical bases of Vessey's theory of cognitive fit (e.g. [Vessey 1991]), and Simon's human information processing theory (e.g. [Simon 1976]). These theories suggest that data quality influences problem enactment both in knowledge (i.e. past attainments) and self-efficacy (i.e. prospective attainments) because:

- Integrated data results in greater task accuracy and faster task completion [Goodhue, Klein and March 2000]. This implies that enhancements in problem enactment with respect to the same task spanning over disperse (i.e. non-integrated) information will be higher because non-integrated data will favor more errors as more complex mental processing will be required (e.g. [Goodhue, Klein and March 2000]).
- Data organized in different ways may provide more or less information for the intended purpose of enacting a problem [Newell and Simon 1972]. This implies that data quality positively influences problem enactment [Seddon 1997].

The link information search behavior → {self-efficacy, knowledge} in Figure 2.1 is supported under the theoretical bases of Kolb's learning theory (e.g. [Kolb 1974]).

This theory suggests that information search behavior influences problem enactment both in knowledge (i.e. past attainments) and self-efficacy (i.e. prospective attainments) because:

- Scanning search favors cognitive integrity [Gonzalez 2001] and it is more likely to lead to mental model building than to mental model maintenance [Vandenbosch and Higgins 1996]. Focused search favors declarative knowledge [Gonzalez 2001] and leads to mental model maintenance [Vandenbosch and Higgins 1996]. Overall, this implies that information search mode positively influences problem enactment and since scanning search favors cognitive integrity, scanning search is more influential than focused search in terms of knowledge enhancements.

2.3.4 Behavioral Factors as the Dependent Variables

The link {data quality, self-efficacy, knowledge} → information search behavior in Figure 2.1 is supported under the theoretical bases of Bandura's SCT [Bandura 1986], and Simon's human processing information theory (e.g. [Simon 1976]). These theories suggest that data quality and problem enactment both in knowledge (i.e. past attainments) and self-efficacy (i.e. prospective attainments) influence information search behavior because:

- Beliefs about capabilities in a given situation-within-context and task-domain knowledge are self-regulatory mechanisms of behaviour [Bandura 1986].
- People tend to approach problems with bounded rationality, involving heuristic information searches (e.g. [Simon 1976]).

The link {self-efficacy, knowledge} → information search behavior in Figure 2.1 is supported under the theoretical bases of Bandura's SCT [Bandura 1986], and Simon's human processing information theory (e.g. [Simon 1976]). These theories suggest that problem enactment both in knowledge (i.e. past attainments) and self-efficacy (i.e. prospective attainments) control information search behavior because:

- Means-ends rationality implies that the 'means' should be appropriate to reach the desired ends (e.g. [Simon 1977]). This implies that the means should be contingent with respect to the ends. Therefore, problem enactment both in knowledge (i.e. past attainments) and self-efficacy (i.e. prospective attainments) as ends influence information search behavior as a means.

- The information that a person chooses to acquire is based on the individual's previous experiences and knowledge (e.g. [Cowan 1986, Lyles and Mitroff 1980, Volkema 1983]). This implies that knowledge in problem enactment (i.e. past attainments) influences information search behavior.
- The stronger the perceived self-efficacy, the higher the goals people set for themselves [Wood and Bandura 1989]. Goals attached to scanning search are considered more challenging than the goals attached to focused search (e.g. [Shaver and Scott 1991]). This implies that strong self-efficacy favors more scanning search and that lower self-efficacy favors focused search.

The link data quality → information search behavior in Figure 2.1 is supported under the theoretical bases of expectancy and adaptive behavior theories (e.g. [Fishbein 1967, Fishbein and Ajzen 1975 and Payne 1976]). These theories support the idea that environmental conditions (e.g. data quality) influence information search behavior because:

- Experimental data suggests that people have an adaptive behavior which adjusts their processing strategies in response to changes in variables such as the number of options available (e.g. [Payne 1976]). Furthermore, no single heuristics does well across all situations; but a person can maintain a reasonably high level of accuracy at a low level of effort by selecting from a repertoire of strategies contingent upon situational demands (e.g. [Payne, Bettman and Johnson 1993, Beach and Mitchell 1978]). This implies that the level of data quality influences the selection of information search heuristics. However, in relation to information search mode, the literature (e.g. [Wetherbe 1991, Rockart and DeLong 1989]) postulates that integrated information is most likely to encourage scanning behavior because integrated information enables executives to explore new relationships and improve their grasp of the business. Unfortunately, empirical research is not conclusive in relation to the information search mode as the dependent variable of data quality. For example, while some research supports the hypothesis that perceived source quality is positively correlated with the frequency of using scanning mode (e.g. [Vandenbosch and Huff 1997, Choo 1993]), Boynton reached an opposing conclusion: the higher (lower) the data quality the less (greater) the intensity in scanning mode [Boynton 1987].

- Data sources of higher perceived quality will be used more frequently than those of lower quality (e.g. [Homburg and Pflesser 2000, Choo 1999, Maltz and Kohli 1996, Menon and Varadarajan 1992, Deshpande and Zaltman 1982, O'Reilly III 1982]). In this case, data source usage means search/request for information in such data source. This implies that the level of data quality influences the intensity of information search behavior in both mode and heuristics.
- Data quality will increase users' satisfaction (e.g. [Gatian 1994]) which will influence their productivity (e.g. [Fishbein 1967, Fishbein and Ajzen 1975]). In other words, it is believed that satisfied users will be more productive. This implies that the level of data quality influences the intensity of information search behavior in both mode and heuristics.

2.3.5 The Environmental Factors as the Dependent Variables

The link {self-efficacy, knowledge, information search behavior} → data quality in Figure 2.1 is supported under the theoretical bases of Bandura's SCT [Bandura 1986], and Kelley's attribution theory [Kelley 1967]. These theories suggest that problem enactment both in knowledge (i.e. past attainments) and self-efficacy (i.e. prospective attainments) and information search behavior influence data quality because:

- As a result of the transactions of information search behaviour and individuals' enhancements on knowledge and beliefs regarding self-efficacy, individuals form a perception of the data quality (e.g. [Doig 2002, Yeoh 2000, Vandenbosch and Higgins 1995]).

The link {self-efficacy, knowledge} → data quality in Figure 2.1 is supported under the theoretical bases of attribution theory [Kelley 1967]. This theory supports that problem enactment both in knowledge (i.e. past attainments) and self-efficacy (i.e. prospective attainments) influences data quality because:

- Individuals tend to consistently attribute causality to personal and/or environmental factors that they think are important (e.g. [Gist and Mitchell 1992, Fletcher, Danilovics, Fernandez, Peterson and Reeder 1986]). This implies that enhancements in problem enactment influence attributions of causality on data quality of such enhancements.

The link information search behavior → data quality in Figure 2.1 is supported under the theoretical bases of Sinkula's organizational learning theory on marketing tasks [Sinkula 1994] and attribution theory [Kelley 1967]. These theories suggest that information search behavior is the main independent variable influencing data quality because:

- Individuals' actions contribute to organizational learning [Huber 1991] by interacting with components of the organizational memory, like a data warehouse (e.g. [Walsh and Ungson 1991]), and by sharing perceptions on such external archives (e.g. [Lee, Courtney and O'Keefe 1992]). This implies that information search behavior performances influence attributions of causality on data quality of such behavior.

2.4 RESEARCH CONTEXT: CUSTOMER RELATIONSHIP MANAGEMENT SUPPORTED BY A DATA WAREHOUSE

2.4.1 Customer Relationship Management

Drawing on the ideas about contexts in management studies (e.g. [Witteloostuijn 1996]) the context of this research is constituted by the marketing processes in an organization, known as customer relationship management (CRM), that are supported by a data warehouse. CRM is the cross-functional business process that drives customer value by the creation and maintenance of business–to-customer durable, close and mutually beneficial relationships (e.g. [Leigh and Marshall 2001]). CRM involves market-sensing and customer-linking activities [Day 1994]. CRM is the label used in the industry for the practice of the currently dominant marketing paradigm known as relationship marketing (e.g. [Morgan and Hunt 1994, Sheth and Parvatiyar 1995, Cravens 1998]) which focuses on customers (e.g. [Payne and Holt 2001]) as opposed to other stakeholders (e.g. employees, suppliers, competitors).

The relationship construct has its conceptual roots in social cognitive research. The literature on interpersonal relationships has experienced phenomenal growth (see [Berscheid 1994] for a comprehensive literature review) and has influenced other disciplines such as marketing (e.g. relationship marketing). The relationship phenomena recognizes that many of the questions addressed by the social and

behavioral sciences directly engage questions about interpersonal relationships [Berscheid 1994]. A relationship is considered a cognitive concept with information about the self, another person and the social interaction between the self and this other person [Berscheid 1994]. Furthermore, the types of information in these cognitive structures are often the expectations about one's own behavior, the behavior of another person, and the nature of the interaction likely to take place between the two [Miller and Turnbull 1986]. The relationship marketing literature extends the relationship construct to agent-to-customer and business-to-customer relationships. Therefore, in contrast to the traditional exchange-oriented model of marketing as a series of potentially unrelated and discrete transactions, relationship marketing is viewed as an ongoing and continuous process of nurturing buyer-seller partnerships over time.

A key construct in social cognitive literature is relationship closeness. Interpersonal relationship closeness is defined as the interdependence of the partner's behaviors, including their emotions and thoughts [Kelley, Berscheid, Christensen, Harvey, Huston, Levinger, McClintoc, Peplau, and Peterson 1983]. Behavior in a close interpersonal relationship is shaped by the situation-within-context characteristics of that relationship [Blumstein and Kollock 1988].

Important constructs in relationship marketing are customer relationship satisfaction, customer trust and customer relationship commitment. Customer relationship satisfaction is defined as a customer's affective state resulting from an overall appraisal of his or her relationship with a firm (e.g. [Anderson and Narus 1990]). Customer trust is defined as a customer's confidence in a firm's reliability and integrity (e.g. [De Wulf, Odekerken-Schroder and Lacobucci 2001]). Finally, customer relationship commitment is defined as a customer's enduring desire to continue a relationship with a firm accompanied by the customer's willingness to make efforts at maintaining it (e.g. [Morgan and Hunt 1994]).

According to Srivastava, Shervani and Fahey [1998], the theoretical foundations supporting the notion that stronger customer relationships are created when the firm uses knowledge about buyer needs and preferences to build long-term relational

bonds are brand equity (e.g. [Keller 1993, Shocker, Srivastava and Ruekert 1994]), customer satisfaction (e.g. [Anderson and Sullivan 1993]), and the management of strategic relationships (e.g. [Anderson and Narus 1996]).

There is also an economic reason in the popularity of CRM in the industry: the well-grounded belief that it is less costly and more profitable to keep current customers than to generate new ones (e.g. [Grönroos 1995, Reichheld and Sasser 1990, Reichheld 1993]). Companies earn a higher return from getting repeat sales from current customers than from spending money to attract new customers [Sheth and Parvatiyar 1995]. From another perspective, Kotler [1988] stated that the seller who knows how to build and manage strong relationships with key customers will have more cross-selling opportunities with their customers. In general there is uncontested and overwhelming evidence in the literature about the benefits of forming customer relationships [Sheth and Parvatiyar 1995].

Technology (e.g. data warehouses) plays a key role in supporting CRM (e.g. [Wang and Spiegel 1994]). In order to truly understand relationships, CRM managers need to integrate many different types of customer information data. For example, some of these data types are general attitudinal judgments (e.g. customer trust and commitment), buying motives/relationship benefits, and psychographic/lifestyle data [Peltier, Schibrowsky and Davis 1998].

Illustrative examples of CRM processes are campaign management and segmentation in the marketing function, lead tracking and contact management in the sales function, and call center support in the customer support function (e.g. [Herschel 2002]).

Kotler [1988] argued that one of the steps in establishing relationship management is to identify the key (top few) customers meriting relationship management. Jackson [1985] argues that relationship management is not effective in all situations but is extremely effective in the right situation. She argued that relationship management investments pay off with customers who have long term horizons, high switching costs and expect consistent and timely service. From my literature review, I conclude

that in CRM processes, each contact with a customer or prospect provides additional information that permits further refinement of the relation-building strategy (e.g. [Gwinner, Gremler and Bitner 1998, Sheth and Parvatiyar 1995]). A single purchase is an event not its ultimate objective (e.g. [Roberts 1996]).

2.4.2 Environmental Limitations

The literature on strategic management (e.g. [Mintzberg 1983, Witteloostuijn 1996, Emery and Trist 1965, Miles and Snow 1978]) has addressed the (organization) environment construct with profusion. According to Witteloostuijn the notion of environment is a 'dustbin' concept, that is, all issues not internal to the organization are captured by the concept of environment [Witteloostuijn 1996]. As such a definition is not manageable, a number of taxonomies have been proposed (e.g. [Mintzberg 1983, Porter 1980]). In this research, I will focus on the task environment, that is the set of immediate stakeholders such as customers and competitors with which the focal organization has to interact directly [Witteloostuijn 1996].

Mintzberg [1983] suggested three dimensions for characterizing environments: complexity, dynamism, and uncertainty. Several authors have identified various sources of uncertainty (e.g. [Van Birgelen, de Ruyter and Wetzels 2000, Lipshitz and Strauss 1997, Orasanu and Connolly 1995]). The most frequently quoted sources of uncertainty in my literature review were: incomplete information, unreliable information, equivocal information, rapidly changing situations, and purposefully misleading information.

Information theory [Shannon and Weaver 1949] uses the concept of entropy as a measure of both the uncertainty and information contained in a message. Entropy establishes that as the information content of a message increases, the level of uncertainty decreases. This normative approach has been extended to perceived uncertainty (e.g. [Brannick 1998]). Brannick states that uncertainty is an expression of an information state and that a positive association is postulated between information and uncertainty [Brannick 1998]. As information increases, perceived environmental uncertainty decreases [Brannick 1998]. Therefore, perceived environmental uncertainty is defined as the absence of information about activities

and events in the environment [Brannick 1998]. The lack of information can manifest itself in two ways: not knowing what information is needed (type A uncertainty), or knowing what information is needed but being unable to access it (type B uncertainty) [Brannick 1998]. From my literature review I conclude that information reduces uncertainty (e.g. [Brannick 1998, Berger 1979]) and that uncertainty on the task environment sometimes leaves no alternative but to make judgements based on experience, a feel of the situation, and a measure of imagination (e.g. [Nevett 1991]).

2.4.3 A Data Warehouse as a Customer Relationship Data Source

A data warehouse is an integrated, non-volatile, collection of unrelated or disparate subject-oriented data sources where each unit of data is relevant to some moment in time and atomic and/or highly summarized (e.g. [Inmon 1996, Marakas 1998, Kelly 1997]). Data stored in data warehouses comes from the many transactional systems that support the mission-critical business processes in the organization and from external data sources (e.g. [Marakas 1998]).

Data warehouses are part of the context of this research. A data warehouse stores data from the task environment. In a way, a data warehouse partially internalizes (e.g. [Witteloostuijn 1996]) the environment. That is why, for the purpose of this research, the quality of data warehouse customer relationship data is considered as an environmental variable.

The fundamental reason for the inclusion of data warehouses in the context of this research is that data warehouses provide a repository of integrated data if they have been properly implemented and maintained. From my literature review there are four streams of research supporting the necessity of data integration. One research stream is the contingency theories in strategic management literature (e.g. [Lawrence and Lorsch 1967]) stating the hypothesis that environmental uncertainty requires integration devices as a mean for differentiation.

The second research stream is the theory of cognitive fit (e.g. [Vessey 1991, Agarwal, Sinha and Tanniru 1996, Goodhue and Thompson 1995]). This stream recognizes the cognitive limitations of a person processing data from disperse sources. The inability

of humans to effectively integrate different types of data has been experimentally documented (e.g. [Slovic, Fischhoff and Lichtenstein 1977, Swinth, Gaumnitz and Rodriguez 1975, Benbasat and Taylor 1982]). The explanation is as follows: Individuals develop a cognitive structure with the information requirements for the task to be performed. If the data required is dispersed across several data sources (i.e. non-integrated), then the individual needs to map (i) his/her initial cognitive structure about the information requirements for the task to (ii) other cognitive structures with the gathered data from each data source. This process is complex and susceptible to errors (e.g. [Goodhue *et al.* 2000, Orasanu and Connolly 1995]).

The third research stream entails theories of sense making (e.g. [Weick 1976]) and its application to fragmented organizations (e.g. [Orton and Weick 1990]). Boynton and Zmud [1987] noted that the successive implementations of information systems have created silos of information. Fragmented historic data favors equivocality [March and Olsen. 1976] and this increases loose coupling in organizations [Weick 1976]. This in turn creates uncertainty and variance in the actions based on such sources of information [Boynton and Zmud 1987]. This variance implies a serious risk for a successful CRM practice because of a lack of consistency (i) interrelating with a customer in the different contact points of a firm (e.g. clerk, phone) and (ii) having a 'single version of the truth' of the customer. On the other hand, combining data from multiple sources enables executives to explore new relationships and improve their grasp of the business (e.g. [Rockart and DeLong 1989]). Research in both marketing and management provides strong support for a positive link between cross-functional integration and firm performance (e.g. [Griffin and Hauser 1996, Song, Xie and Dyer 2000]). Wetherbe [1991] convincingly linked cross-functional integration with information integration.

The fourth research stream includes theories on organizational learning where organizational memory is information stored from an organization's history that can be recalled as needed (e.g.[Walsh and Ungson 1991, Daft and Weick 1984]). Databases (e.g. data warehouses) are artifacts implementing organizational memory (e.g. [Huber 1991]).

It would be unfair to neglect the relevant role of qualitative analysis (e.g. focus groups) in marketing in general, and CRM in particular, for gathering evidence. However, the dominant approach to understanding knowledge/information utilization in marketing theory is confined to the scientific/quasi scientific type of information to the exclusion of the narrative/subjective mode (e.g. [Brannick 2000]) and data warehouses play a key resource providing data (e.g. [Hair Jr., Babin, Money and Samouel 2003]) in marketing information systems (e.g. [Fletcher 1995]).

The view of market analysts (e.g. [Peynot and Kinikin 2003]) is that a proper implementation of a data warehouse in a CRM context will provide the data foundation for enabling CRM applications such as fulfillment, customer segmentation, targeted marketing, cross-selling, customer loyalty (i.e. attrition), profitability analysis, integrated billing, market basket analysis and other forms of combining customer touchpoints with external data to generate a valuable business asset. The following are some examples of the enabling role that a data warehouse has in a CRM context:

- The data warehouse enables the storage of huge data volumes from, for example, electronic points of sales systems, automatic teller machines, call detail records, and click streams in Internet accesses.
- The data warehouse provides a single view of a customer profile including demographics and lifestyle data by integrating the transactional and contact data history about this customer. This single view of the customer helps in managing the relationship with the customer. Cross selling and customer retention programs are enabled by such integrated information.
- The data warehouse enables the creation of new customer profile information by linking the profiles across customers like householding.
- The data warehouse improves understanding of the history of the customer relationship by storing high volumes of detailed data about the transactions of such customers.
- The data warehouse enables the creation of new information at the market level by storing acquired external data, for example lists of prospects, and by identifying target markets to be addressed on a one-to-one basis.

- The data warehouse enables event-driven initiatives by using triggers, user defined functions and engines.
- The data warehouse will enable lifetime value recognition by providing historic data to forecasting models and comparing it with existing validated patterns.

However, as systematically reported in market studies (e.g. [Peynot and Kinikin 2003]), one of the biggest challenges to CRM success is the poor quality of customer relationship data. From my literature review, I conclude that data warehouses are a key enabler of data integration.

2.5 RESEARCH SITUATION: MAKING SENSE OF PROBLEMS

2.5.1 Problems

The term problem refers to the individual's perception of a variance, or a gap, between the present and some desired state of affairs (e.g. [Simon 1977, Smith 1990]). Something that is difficult to achieve due to the lack of awareness on how to manage the situation or the lack of resources, both fit in the previous definition as the present state of affairs.

Dery [1983] complains about the tendency in the literature which presumes that individuals are facing predefined problems – those stated independently of the individual's perceptions. He further argues that problems do not present themselves as structured or ill-structured, nor do they come as complex or simple problems. Structured problems are structured because we choose to treat them as such. Such critique is important for this research because it highlights the relevance of defining a problem. For the purpose of this research, the term problem statement refers to the formulation in linguistic terms of the problem elements and its structure (e.g.[Smith 1989, Abualsamh *et al.* 1990, Pitz *et al.* 1980]). Some theorists regard problem statements as the best defense against a type III error, that is, addressing and potentially investing resources in the wrong problem [Smith 1989]. Likewise, a problem statement will be determinate in the future course of action (e.g. [Volkema 1983, Dery 1983, Winkler 1982, Mintzberg, Raisinghani and Théorêt 1976, Mitroff and Featheringham 1974, Abualsamh *et al.* 1990]).

This constructivist approach to problems implies that what is perceived by one individual as a problem may not be viewed as a problem by another member of the same organization [Smith 1990]. This disparity may occur because the problem doesn't really exist according to the collective opinion of the social environment of the individual or simply because the problem really does exist and there is no consensus in the social environment on a common formulation.

Most of the early literature on problem-solving (e.g. [Simon 1960]), as Dery [1983] criticizes, addresses the structural aspect of the problem statement as an intrinsic property of the problem and not as a result of a cognitive process. The conclusion from my literature review is that assessments like well-structured problems versus ill-structured problems are valuable frames that can be applied post-hoc to the cognitive process of enactment. For example, when Orasanu and Connolly [1995] say that situations in natural settings are characterized by ill-structured problems, we can adjust this to "problems formulated in natural settings are characterized by their ill-structureness".

One aspect in the definition of problem statements is its neutrality (i) with respect to any framing, and (ii) with respect to any solution approach. A problem statement should not include an interpretation of the problem in terms of a symbolic label. Examples of frames are customer oriented (e.g. [Day and Nedungadi 1994]), drama (e.g. [Corey and Wilson 1994, Burke 1969]), negative-losses (e.g. [Dutton and Jackson 1987, Tversky and Kahneman 1981]), threats and opportunities (e.g. [Kotler 1988]), P problems (i.e. deterministic solution) and NP problems (i.e. hard to verify a-priori if there is a solution). Likewise, the problem statement does not preclude any solution at all; in fact, the problem might by an NP problem.

Massey and O'Keefe [1993] validated Sakman's multi-attribute model of problem definition quality [Sakman 1985] and concluded that comprehensiveness, which covers Sakman's structure and other attributes, was the key characteristic in problem statements. Comprenhensiveness has been studied in the cognitive literature under the label of integrity complexity. Integrative complexity refers to the level of comprehensiveness -i.e. number of factors in the cognitive structure- and

connectedness -i.e. links among the factors in the cognitive structure- (e.g. [Sullivan and Weaver 2000, Wang and Chan 1995, Feist 1994]).

Orasanu and Connolly [1995], talking about the task environment in natural decision making situations, says that natural settings are characterized by (i) uncertainty, (ii) dynamicity, (iii) shifting, ill-defined, or competing goals, (iv) action/feedback loops, (v) time stress, (vi) high stakes, (vii) multiple players, and (viii) ill-structure problems. From my literature review, I conclude that the task environment creates more obstacles than providing help for the important task of formulating problems. In general, making sense of some of the problems facing organizations is not easy (e.g. [Volkema 1983]).

2.5.2 Cognitive Processes

Problem statements are cognitively construed. Cognitive processes refer to the mental processes involved in the acquisition, organization and use of information [Bandura 1994]. Enactment, learning, self-regulation and bounded rationality are the key cognitive processes that play a role in the situation that this research focuses on.

Enactment is the first cognitive process in my selection of relevant cognitive processes in this research. Enactment is a genuine contribution of the sense making theory (e.g. [Weick 2000]). Sense making is a high order cognitive process intended to reduce equivocality, or multiple meanings, in the information [Weick 2000]. For example, making sense of a customer relationship problem means that heedful interrelating connects sufficient individual knowledge with situational demands [Weick 1993].

From my literature review, I concluded that the terms understand, make sense and comprehend refer to the same concept. Enactment is a sense making process (e.g. [Weick 2000]). By enactment it is meant the generation of information, plausible interpretations of a (problematic) situation, and actions to be realized (e.g. [Weick 2000]). Terms like "problem statement" and "problem focus" [Kuhlthau 1993] refer to the concept of enacted problem.

Weick lists seven attributes of sense making [Weick 1995]: (1) it is grounded in identity construction, (2) it is retrospective, (3) it is enactive of sensible environments, (4) it is social, (5) it is ongoing, (6) it is focused on extracted cues, and (7) it is based in plausibility rather than accuracy. The basic idea of sense making is that reality is enacted in an ongoing accomplishment that emerges from efforts to create order and make retrospective sense of what occurs [Weick 1993].

Because sense making is grounded in identity construction and based on plausibility, we might expect differences owing to differences in the individual [Taylor 1999]. For example, if one CRM manager says that the problem in a given customer relationship is low customer satisfaction and another says it is competitive pressure, marketing senior executives might assume one of the CRM managers was correct, and disregard the other; other senior executives would look for another third problem statement, or explore the reasons for their problem formulations in more depth. But very few senior executives would take an average (i.e. settle for an inconclusive assignment of probabilities to the two possibilities) (e.g. [Cohen 1995]).

Because sense making is enactive of sensible environments and focused on extracted cues, we might expect individuals in different environments to make sense of things differently [Taylor 1999]. Identity construction, plausibility, and a particular set of extracted cues also seem tied to the characteristics of a particular job and we might expect people with different jobs to make sense of data differently [Taylor 1999].

Learning is the second cognitive process in my selection of relevant cognitive processes in this research. There are a number of learning theories (e.g. double loop learning theory [Argyris 1976], experiential learning [Kolb 1984]). They stress acquisition of knowledge and cognitive structures and the processing of information and beliefs. Learning refers to the sequence of information processing activities involved in the creation of knowledge [Day 1994]. Furthermore, some authors believe that behavioral change is required for learning (e.g. [Fiol and Lyles 1985, Levitt and March 1988, Huber 1991, Argyris and Schön 1978]) while others believe that new ways of thinking are enough [De Geus 1988]. Some authors stress open-minded

approaches to problem solving [Senge 1992] which include new ways of formulating problem statements.

From my literature review I have concluded that learning theories and sense making theory differ in the following: (i) Knowledge is an output of learning while in sense making it is an input, and (ii) information is an output of sense making while in learning it is an input. In simple terms, enactment will allow formulating a problem statement and behaving accordingly and learning will allow acquiring knowledge about the problem statement and behavior.

Self-regulation is the third cognitive process in my selection of relevant cognitive processes in this research. Self-regulation is based on self-beliefs. What do people believe about their cognitive capabilities (e.g. enacting CRM problems)? People tend to make attributions of cognitive capability (i.e. self-beliefs on their cognitive capabilities) either as an acquirable skill or as a stable entity (i.e. fixed inherent capacity) [Dweck and Elliot 1983]. Those who view a cognitive ability (e.g. enacting CRM problems) as an acquirable skill regard it as continually enhanceable through knowledge and the perfection of one's competencies. They adopt an inquiring learning goal. They seek challenges that provide opportunities to expand their knowledge and competencies. For them, errors are regarded as a natural, instructive part of an acquisition process. They judge their capabilities more in terms of personal improvement than by comparison against the achievement of others. People with self-beliefs in their cognitive ability (e.g. enacting CRM problems) as a more or less fixed capacity regard their performance level as a diagnostic of basic intellectual aptitude. For them, errors and deficient performances, therefore, carry personal and social evaluative threats [Wood and Bandura 1989].

Bandura is the most prominent among recent voices calling for a new perspective in self-beliefs. Bandura identified that individuals create and develop self-perceptions of capability [Bandura 1977]. He termed such self-beliefs as self-efficacy, that is, the self-belief in one's capabilities to organize and execute the sources of action required to manage prospective situations [Bandura 1977]. Later in [Bandura 1986] he provided a theory of human functioning, SCT, that identifies the central role of self-

efficacy in self-regulation. This theory argues that human functioning is the result of a determinism of personal, behavioral, and environmental factors and that self-efficacy becomes instrumental to the goals that individuals pursue and to the control individuals are able to exercise over their environments.

Bounded rationality is the fourth cognitive process in my selection of relevant cognitive processes in this research. Nobel prize-winning scholar Herbert A. Simon stated in his seminal work [Simon 1960] and other subsequent works [Simon 1962, Simon 1965, Simon and Newell 1971, Newell and Simon 1972, Simon 1976] the notion of rationality in decision making and problem solving situations. Rationality in a problem enactment situation is the extent to which the sense making process involves the collection of information relevant to the problem, and the reliance upon analysis of this information in enacting it (e.g. [Simon 1978]). Therefore, rationality involves behavior, that is information search, and cognitive processes of analysis and decision making. At the same time, Simon theoretically recognized cognitive constraints to rationality in creating complex cognitive structures and assimilating large amounts of information. Many studies have confirmed such limitations, for example, the amount of information that a person can assimilate (e.g. [Miller 1956]), integration of data from disperse data sources (e.g. [Goodhue and Thompson 1995]), and the development of complex cognitive structures (e.g. [Pratch and Jacobowitz 1998, Suedfeld, Tetlock and Streufert 1992]).

Simon formulated a theory on rational human information processing contending that as a result of uncertainty and cognitive constraints to rationality, people tend to approach problems with bounded rationality, involving heuristic information searches, means-ends analysis and the use of 'satisficing' criteria to make choices. The term heuristic refers to simplifying routines used by people in their information processing activities in order to search and filter information coping with their cognitive limitations [Simon 1976]. People rely on a limited number of heuristics (e.g. [Tversky and Kahneman 1974, Hogarth and Makridakis 1981, Schwenk 1984, Hogarth 1987]). Selection or rejection of information is influenced by the individual's preferred heuristics (e.g. [Tversky and Kahneman 1974, Hogarth and Makridakis 1981, Schwenk 1984, Hogarth 1987]). These heuristics are two-edged, for while they

reduce mental effort in sense making and are highly selective (i.e. explore only a minuscule fraction of the total available information) [Simon 1977], their use can also lead to systemic biases or errors in judgment (e.g. [Tversky and Kahneman 1974]).

When properly applied, information search heuristics reduce search time. The ability to quickly access and effectively process data is the essence of domain-specific expertise (e.g. [Newell and Simon 1972]). In general, the literature on managerial competences considers information search as a key competence (e.g. [Schroder 1989, Spencer and Spencer 1993]).

Means-ends analysis follows the causal determinacy principle by which 'means' (e.g. resources, time, choices, actions) should be appropriate to reach the desired ends [Simon 1977]. To judge a behavior as rational is to be able to say that the behavior is understandable within a given situation-within-context of reference. That behavior may, however, appear as rational to the actors in a situation but irrational to an observer. Behavior inconsistent with the actor's frame of reference may be deemed irrational, as may behavior for which an observer can find no explanation [Butler 1996]. Reed described contextual rationality as action motivated to create and maintain institutions and traditions that express some conception of right behavior and a good life with others [Reed 1991]. Contextual rationality is sensitive to the fact that social actors need to create and maintain socially accepted norms that sustain and enrich their relationships. Thus, organizations and data sources, like data warehouses, become important because they can provide meaning and order in the face of environments that impose ill-defined, contradictory demands (e.g. [Weick 1993]).

From my literature review I conclude that individuals can be good at decision making and still falter because of deficient problem enactment. The world of decision making is about strategic rationality. It is built from clear questions and clear answers that attempt to remove ignorance (e.g. [Daft and Macintosh 1981]). The world of problem enactment is different. Problem enactment is about contextual rationality. It is built out of vague questions, and muddy answers and the cognitive processes of enactment,

learning, self-regulation and bounded rationality play a key role in reducing equivocality.

2.5.3 Information Search Behavior

From my literature review emerges a view of information search as a rational behavior in which a person is actively constructing a new understanding from the information encountered (e.g. [Kuhlthau 1999]) where such behavior is characterized by the way the individual's choices link successive searches [Dervin and Nilan 1986]. When should one stop searching for information? Kuhlthau [1999] introduced the concept of "enough". The application of her notion implies that "enough" is reached when the requirements of information, represented by an 'empty' cognitive structure, for formulating a problem statement are met. This notion of gap is also suggested in the anomalous states of knowledge theory (e.g. [Belkin 1980]) and labeled as an anomalous state of knowledge.

A second perspective of "enough" is related to the uniqueness of information during the information search process for formulating a problem statement. At the beginning of an information search process we can expect that the likelihood of encountering uniqueness (new information) will be high and redundancy (familiar information) to be low. As the process progresses and the person learns more about the problem there is likely to be more of a balance between the two types of information. Finally, at the close of the process we can expect that the ratio may be reversed with uniqueness low and redundancy high. Therefore, uncertainty may be associated with high uniqueness and "enough" with high redundancy (e.g. [Kuhlthau 1999]).

A third perspective of "enough" is related to the expected incremental value of increasing the quality of data (e.g. missing data, wrong data) with respect to the resources consumption (e.g. time, money) allocated [Marakas 1998]. A fourth perspective of "enough" is provided by the exit criteria of the information search heuristic being used. The term exit is borrowed from the information processing theory (e.g. [Miller, Galanter and Pribram 1960]) to express the satisfying criteria for ending an information search.

Considerable empirical research has defined rationality as the collection and analysis of information (e.g. [Mintzberg *et al.* 1976, Bourgeois III and Eisenhardt 1988, Fredrickson1984]) and several studies have found broad variation in rationality, that is, the degree of information collection and reliance upon analysis (e.g. [Dean Jr. and Sharfman 1993]). On the topic of decision making, Jones and McLeod Jr. [1986] contend that little research has examined variation among decision makers regarding where and how information used in decision making is obtained. Finally, with regard to research on information search, in recent years there has been a lack of theory-testing field work [Huber 1991]. The term behavior refers to the way in which an individual acts or works [Oxford 1993]. There are two not mutually exclusive types of information search behaviors that I am interested in studying in customer relationship problem sense making, which is the type of situation that this research focuses on. One type is information search mode and the other type is information search heuristics. Both may be exhibited by a person in parallel.

Scanning search and focused search are two information search modes studied in the literature. I am adopting the term mode from Churchman's philosophically based inquiring modes [Churchman 1971]. Scanning search mode is the proactive behavior people exhibit when they browse through information without a particular problem to solve [Aguilar 1967]. Scanning is typically done when executives explore emerging trends, changes, opportunities, and issues to evaluate how they will impact corporate decisions (e.g. [Sawyerr, Edbrahimi and Thibodeaux 2000]). Scanning also involves dividing the environment into meaningful sectors, collecting data, and forecasting changes in key variables (e.g. [Sawyerr *et al.* 2000]). The literature is not conclusive in that scanning is not necessarily characterized by problem search but certainly it is frequently suggested that this is the case (e.g. [Stubbart 1989, Daft and Lengel 1984, Kiesler and Sproull 1982]). Aguilar [1967] found that customers and competitors received the greatest scanning attention by executives.

Focused search mode is the reactive behavior people exhibit when they are looking for information specific to a problem to be addressed or question to be answered [Huber 1991]. Focused search is the typical information search initiated by a business question (e.g. [Vandenbosch and Ginzberg 1997]). For the purpose of this research, I

define the concept of business question (e.g. [Schank 1988]) as a sentence worded or expressed in such a form so as to elicit information. The information technology literature addresses business questions through queries (e.g. [Vandenbosch and Ginzberg 1997]).

In general, the goals attached to scanning search (e.g. emerging trends, changes, opportunities, and issues) are considered more challenging than the goals attached to focused search (i.e. respond to a business question), in the sense that scanning search is considered a riskier behavior than focused search from the perspective of goal attainments (e.g. [Shaver and Scott 1991]).

Representativeness, availability, anchor and adjustment, and positivy are four information search heuristics studied in the literature. There is a substantial theoretical foundation on the development and use of information-processing heuristics although most of it has been on situations where the use of heuristics may bias decision processes.

The representativeness search heuristic [Kahneman and Tversky 1972] refers (i) to assessing the probability of a situation as representative of a category [Wright 1980], or (ii) to making generalizations based on new information about a sample [Wright 1980] (e.g. managers may quickly categorize a customer as representative of a segment). This heuristic implies that the search ends when there is a satisfying fit between information about a situation and information about a category (e.g. [Chi and Fan 1997]). This heuristic is based on the similarity to previously accepted patterns [Averett 1991].

People use search heuristics of representativeness when they judge the similarity of the gathered information to profiles that they believe to be representative of a category. Representativeness is a search heuristic method that involves search and compare. When looking for information to enact an unfamiliar problem, a person who employs this heuristic will search for information on more familiar problems and select the closest to the unfamiliar problem. Then, the information on this selected problem statement is adopted for formulating the new problem.

The availability search heuristic [Tversky and Kahneman 1972] refers to assessing the probability of a situation as a function of prior situations [Wright 1980] (e.g. a marketer considering a series of occurences of actual costs incurred in past editions of a campaign when estimating the cost for a new edition of such a campaign). This search heuristic implies a search for recent, salient, easily accessible information about relevant precedents. The search ends once recent, salient information about a relevant precedent is found (e.g. [Chi and Fan 1997]). This search heuristic is based on the availability of information about situations at a critical time [Averett 1991]. People use search heuristics of availability when they recall familiar instances. When looking for information to enact a new instance of a familiar problem, a person who employs this heuristic will search for information used in the formulation of the most recent problem statement of such type.

Anchoring and adjustment search heuristic [Tversky and Kahneman 1974] refers essentially to the trial and error method [Chi and Fan 1997] (e.g. a marketer setting the price of a product starting with a baseline price and making a number of impact analyses in several of the cost components). This search heuristic implies a recursive process and each step involves a search for additional information and an adjustment of the previous assessment. The search ends once the adjustments are not improving an implicit/explicit value function on the information found (e.g. [Chi and Fan 1997]). Anchoring and adjustment search heuristic simply means that individuals start at one place from that initial point. Individuals try to get "close" and then make adjusts from an initial point by obtaining and using additional information.

People use the search heuristic of anchoring and adjustment when they are trying to formulate a problem statement that is plausible with respect to more than one criteria. They do so by starting from an initial problem statement (the anchor) and adjusting it by adding and dropping information to arrive at a final formulation.

Positivy search heuristic [Evans 1989] refers to confirming the probability of a situation. This heuristic implies a search for information that is fundamentally consistent with existing beliefs, theories and cognition. The search ends once the

information found confirms the probability of a situation [Evans 1989]. People use search heuristic of positivy when they are trying to confirm the plausibility of an already formulated problem statement with enough subjectively considered evidence.

From my literature review I conclude that rational information search behavior in problem sense making is exhibited as a combination of information search modes (i.e. scanning, focused) and information search heuristics (i.e. representativeness, availability, anchor and adjustment, and positivy).

2.6 CUSTOMER RELATIONSHIP DATA QUALITY FOR PROBLEM ENACTMENT

Customer relationship data quality for problem enactment is part of the situation that this research focuses on. From my literature review, four research streams emerge that are related to this construct (see Table 2-1).

The first stream of research, which I will label as generalist, is characterized by neglecting the relevance of situational and contextual aspects. Quality of data is assessed in terms of inherent attributes like integration, completeness (i.e. missing data), accuracy, timeliness, richness, that is the extent to which data carries information, equivocality (i.e. multiple meanings), and trustworthiness (see Table 2-2).

This stream evolved first to a structured approach by recognizing the high order of the data quality construct, that is its dimensional nature. Zmud was the first to probe the dimensionality of data quality [Zmud 1978].

In general, the dimensional aspects of data are overwhelmingly accepted (e.g. [O'Reilly III 1982, DeLone and McLean 1992, Huang, Lee and Wang 1998, Wang, Reddy and Kon 1995, Fox, Levitin and Redman 1994, Jarke, Jeusfeld, Quix and Vassiliadis 1999, Wang and Strong 1996, Eppler and Wittig 2000, Fedorowicz and Lee 1998, Wang, Storey and Firth 1995, Wand and Wang 1996, Kim 1989]). In general, data quality is considered a first order construct. Some of this research

includes a contextual dimension but it is worded in generic terms with the purpose of being applied in any context (e.g. [Huang *et al.* 1998]).

QUALITY OF DATA RESEARCH STREAMS	SUMMARY
Generalist	Quality of data is assessed in terms of inherent attributes recognizing the high order of the data quality construct (i.e. its dimensional nature).
Contextually utilitarian	Quality of data is assessed in terms of its utility in a given context. It is the independent variable and value is the dependent variable.
Situational-within-context utilitarian	Quality of data is assessed in terms of its utility in a given situation within a context.
Probabilistic-algorithmic	Quality of data is assessed in terms of either classic Bayesian probability theory or extensions of this theory. Uncertainty and quality of data are complementary viewes.

Table 2-1 Research streams related to quality of data

A further development on this construct was the influence of TQM ideas (e.g. [Ahire, Golhar and Waller 1996]) and the replication/adoption of process oriented ideas to the construct (e.g. [English 1999, Huang *et al.* 1998]). These works tend to be descriptive and their contribution is very popular among practitioners although the validity of the concepts remains untested for most of its principles. For example, Wang considers data as a manufactured product and suggests aninput-process-output approach to data quality [Wang *et al.* 1995]. Another data quality approach is the data life cycle which focuses on the sequence of activities from creation to disposition of data [Redman 1992]. In general, from the myriad dimensions that this research stream suggests about data quality, I conclude that integration, richness, trustworthiness, and equivocality are key dimensions of data quality for problem enactment (see Table 2-2).

The second stream of research, that I will label as contextually utilitarian (see Table 2-1), is based on the concept of value of the available information in a given context

(e.g. [Low and Mohr 2001, Moorman and Austin 1995, Vandenbosch and Higgins 1995]). The term value is used in some literature as a higher order construct than "utility" (e.g. [Doig 2002]), although the terms "value" and "utility" are frequently used for the same concept (e.g. [Swartzmeyer 1987]).

QUALITY OF DATA DIMENSIONS	REFERENCES IN THE GENERALIST RESEARCH STREAM
Integration	(e.g. [Doig 2002, Goodhue *et al.* 2000, Haley 1997, Goodhue Wybo and Kirsch 1992, Peltier *et al.* 1998, Wetherbe 1991, March and Olsen 1976, Codd 1970])
Completeness	(e.g. [Fedorowicz and Lee 1998, O'Reilly III 1982, Boynton and Zmud 1987, Orasanu and Connolly 1995, Moenaert and Souder 1996, Wang and Strong 1996, Huang *et al.* 1998, Rudra and Yeo 1999])
Accuracy	(e.g. [Wang and Strong 1996, Huang *et al.* 1998, Rudra and Yeo 1999, Swanson 1987, Haley 1997])
Timeliness	(e.g. [Song *et al.* 2000, Moenaert and Souder 1996])
Richness	(e.g. [Weick 2000, Moorman 1995, Daft and Lengel 1984])
Equivocality	(e.g. [Goodhue *et al.* 2000, Jarke *et al.* 1999, Franz 1999, Brannick 1998, Swanson 1987, Daft and Lengel 1986, Daft and Weick 1984, Daft and Macintosh 1981, Weick 1979])
Trustworthiness	(e.g. [Grooms 2001, Jarke *et al.* 1999, Moenaert and Souder 1996, Moorman, Zaltman and Deshpande 1992, Wixom and Watson 2001, Song *et al.* 2000, Swanson 1987, Srinivasan 1985, Seddon and Kiew 1994, O'Reilly III 1982, Kettinger and Lee 1994, Venkatesh and Davis 2000, Choo 1993])

Table 2-2 Literature addressing the dimensionality of data quality

This stream of research constructs data quality as a different concept with no utilitarian aspect at all. Here, data should be considered as an asset of the organization (e.g. [Glazer 1991, Davenport, Eccles and Prusak 1992]) and in general, the

contextually utilitarian approach contends that data quality is the independent variable and value is the dependent variable.

Several works argue (e.g. [Glazer 1993]) that measuring data assets is a primary mechanism by which an organization becomes an information-intensive organization. This premise is generally found in the literature on information economics (e.g. [Laffont 1989, McDonough 1963, Parker, Benson and Trainor 1988]) and information systems evaluation (e.g. [Remenyi, Sherwood-Smith and White 1997]).

This research stream has provided different dimensions to data quality that are contextually significant like value, relevance and usefulness. However, in the literature on the marketing information processing (e.g. [Sinkula 1994, Chapman 1989]), as expected, there is more specificity. This literature adopted the notion of insights from the problem solving literature (e.g. [Kaplan and Simon 1990, Davidson, Deuser and Sternberg 1994]). The concept of data insights refers to the attribution of some data's cognitive utility (e.g. [Kaplan and Simon 1990]). This term is used in the marketing information processing literature for customers (customer insights) to designate the data that is relevant for deriving information on customer profile and future customer behavior.

Insightful customer data is based on customer attitudes, that is, learned predispositions to respond in a consistently favorable or unfavorable manner with respect to a given subject, characteristic and needs (e.g. [Fishbein and Ajzen. 1975]). A favorable attitude should be considered a strong predictor of costumer-behaviour intentions. Examples of customer insights are propensity to buy and propensity to churn.

Similarly, insightful competitor data is based on the strengths, weaknesses, capabilities and strategies of competitors. It typically contains descriptive secondary data from market analysts and primary data from transactions of competitor's customers or direct observation (e.g. [Porter 1980, Prescott 1995]). Examples of competitor insights are product price and product sales per area. See Table 2-3 for a comprenhesive summary of references on this research stream.

QUALITY OF DATA CONTEXTUAL UTILITY	REFERENCES IN THE CONTEXTUALLY UTILITARIAN RESEARCH STREAM
Value	(e.g. [Doig 2002, Vandenbosch and Higgins 1995, King and Epstein 1983, Munro and Davis 1977, Gallagher 1974, Swanson 1987])
Relevance	(e.g. [Franz 1999, O'Reilly III 1982, Zmud 1978])
Usefulness	(e.g. [Mahmood and Medewitz 1985, Lucas Jr. 1981])
Customer insights	(e.g. [Davenport, Harris and Kohli 2001, Doig 2002, Day 1994, Narver and Slater 1990])
Competitior insights	(e.g. [Doig 2002, Beal 2000, Day 1994, Narver and Slater 1990])

Table 2-3 Literature addressing quality of data contextual utility

Likewise, the information systems literature has adopted the notion of quality as part of higher order constructs (e.g. [Kettinger and Lee 1994]) with a strong emphasis on the idea that perceived quality is thought to be cognitive (e.g. [Roest and Pieters 1997]) and therefore, being enacting problems the task at hand, data quality should have such utility as to enable the advocated fit by this stream of research.

In general, from my literature review on this research stream, I conclude that customer data insights and competitor data insights are key dimensions of data quality for problem enactment.

The third stream of research, which I will label as situational-within-context utilitarian (see Table 2-1), is based on the theory of cognitive fit (e.g. [Goodhue and Thompson 1995, Goodhue *et al.* 2000, Huang *et al.* 1998]). The theory of cognitive fit posits that synergistic problem solving will take place when there is a match between the different problem information requirements and the information used. This means that superior problem enactment performance will result when the problem information needs and the information gathered fit (e.g. [Agarwal *et al.*

1996, Vessey 1991]). The term cognitive fit refers to the match between a task's cognitive structure demands and the information used (e.g. [Vessey 1991, Agarwal *et al.* 1996, Goodhue and Thompson 1995]). See Table 2-4 for a comprenhensive summary of references on this research stream.

QUALITY OF DATA SITUATIONAL UTILITY	REFERENCES IN THE SITUATIONAL-WITHIN-CONTEXT UTILITARIAN RESEARCH STREAM
Research on cognitive fit	(e.g. [Fedorowicz and Lee 1998, Bailey and Pearson 1983, Venkatesh and Davis 2000, Moenaert and Souder 1996, Jenkins and Ricketts 1985, Wang and Strong 1996, O'Reilly III 1982, Huang *et al.* 1998, Swanson 1987, Haley 1997, Streufert 1973, Boynton 1987, King and Epstein 1983])

Table 2-4 Literature addressing quality of data situational utility

In general, from my literature review on this research, I conclude that problem enactment utility, that is, the data is helpful for enacting customer relationship problems, is a key dimension of data quality for problem enactment.

The fourth stream of research, which I will label as probabilistic-algorithmic (see Table 2-1), is based on either classic Bayesian probability theory or extensions of this theory. The following data quality issues are addressed in this type of research: missing data, data errors, missing attributes, and missing metadata. Classic probability theory has been successfully applied in the treatment of missing data (e.g. [Little and Rubin 2002, Agarwal and Parthasarathy 2001]) and it is very popular in empirical studies in social sciences (e.g. surveys). Information theory [Shannon and Weaver 1949] builds on classic probability theory in the concept of entropy as a measure of both the uncertainty and information contained in a message.

The treatment of data errors can be broadly divided into systematic errors and non-systematic errors. Systematic errors can sometimes be detected and corrected (e.g. [Brazdil and Clark 1990]). Non-systematic errors, usually called noise, are much

harder to detect and correct. Several techniques have been devised to cope with noisy data (e.g. [Quinlan 1989, Bratko 1994]). In general, fuzzy logic, evidence theory, and rough set theory are three theoretical approaches for addressing data errors. In fuzzy logic the truth value of a proposition can take on any subjective value between 0 and 1 (e.g. [Shan, Ziarko, Hamilton and Cercone 1996]). Evidence theory is a generalization of classic (Bayesian) probability theory. In classic probability theory an objective/subjective probability function is applied to each individual proposition, whereas in evidence theory a subjective belief function is assigned only to the propositions that are supported by some evidence. Belief in a proposition does not imply a complementary belief in its negation (e.g. [Bell 1993]). Rough set theory represents a set by a lower approximation and an upper approximation (e.g. [Pawlak, Grzymala-Busse, Slowinski and Ziarko 1995]). Elements belonging to the lower approximation of a set are definitely elements of the set. The boundary region of a set consists of elements from the upper approximation of the set that are not members of its approximation. Elements belonging to the boundary region of a set are possibly members of the set. In essence, the management of uncertainty is done through the computation of an objective quality measure for the lower and upper approximations of a set.

Missing attributes and irrelevant attributes can be detected with algorithms (e.g. [Uthurusamy, Fayyad and Spangler 1991]). Inference of metadata (e.g. a logical data model) can be done from actual data or functional dependencies using data reverse engineering techniques, for example rule induction (e.g. [Chiang, Barron and Storey 1994, Hainaut, Henrard, Roland, Englebert and Hick 1996, Simoudis, Livezey and Kerber 1995, Piatetsky-Shapiro and Frawley 1991]).

From my literature review on this stream of research, I conclude that uncertainty is a complementary aspect of data quality for customer relationship problem enactment.

Overall, I conclude that (i) the four streams of research are active, (ii) data quality is dimensional, (iii) in the context of marketing activities, users' perceptions on information value are essentially interpreted in terms of instrumental utility; perceptions of low quality have major impact in the perceived (low) value of such

data (e.g. [Doig 2002]); and there are expectations of a direct link between high data quality and perceived high value (e.g. [Doig 2002]), and (iv) data quality defined in situational-within-context utilitarian terms is the most rigorous approach (see Table 2-5).

Therefore, drawing on the previous findings and conclusions I define the quality of data warehouse customer relationship data for problem enactment as the manager's perception on the extent to which the information derived from the available customer relationship data in the data warehouse cognitively fits with the manager's CRM cognitive structures for problem enactment.

QUALITY OF DATA LITERATURE REVIEW SUMMARY	REFERENCES
Quality of data is dimensional	(e.g. [O'Reilly III 1982, DeLone and McLean 1992, Huang *et al.* 1998, Wang *et al.* 1995, Fox *et al.* 1994, Jarke *et al.* 1999, Wang and Strong 1996, Eppler and Wittig 2000, Fedorowicz and Lee 1998, Wang *et al.* 1995, Wand and Wang 1996, Kim 1989])
Quality of data has instrumental utility	(e.g. [Doig 2002])
There is a direct link between high data quality and perceived high value	(e.g. [Doig 2002])
Conceptual focus in quality of data is reached in situational-within-context utilitarian terms	(e.g. [Goodhue and Thompson 1995, Goodhue *et al.* 2000, Huang *et al.* 1998])

Table 2-5 Key findings in the literature about quality of data

Furthermore, considering what the literature says about uncertainty (i) in terms of incomplete information, unreliable information, equivocal information, rapidly changing situations, and purposefully misleading information (e.g. [Van Birgelen *et*

al. 2000, Lipshitz and Strauss 1997, Orasanu and Connolly 1995]), and (ii)in terms of its relation to information, that is, as information increases uncertainty decreases (e.g. [Brannick 1998]), I conclude that my suggested definition of data quality is the reverse construct of perceived uncertainty enacting customer relationship problems. That is, perfect quality means zero perceived customer relationship problem enactment uncertainty and vice versa.

2.7 SUMMARY OF LITERATURE REVIEW

In this chapter, literature on data quality, cognitive factors, human information processing/problem solving, marketing information systems, and behavior have been reviewed from the common perspective of data quality. SCT is the overall theoretical body supporting a triadic causal association among cognitive, behavioral and environmental variables. Every dyadic causal association is supported by specific literature addressing the constructs of interest in this research, that is, data quality, problem enactment and information search behavior. This review of additional literatures was needed in order to provide a robust theoretical foundation for the hypotheses that will be postulated in the next chapter. So, for example, while the SCT supports the general link {cognitive, environmental} → behavior, I deemed advisable to also have theoretical support to its specific instantiation {data quality, self-efficacy, knowledge} → information search behavior. This is also done for every direction in the three postulated associations. Overall, from my literature review, I have concluded that the SCT holistically explains the links postulated by all the other reviewed literatures.

3. CONCEPTUALIZATION

3.1 INTRODUCTION

This chapter starts with the CONTEXT, SITUATION AND UNIT OF ANALYSIS RESEARCH SPECIFICATIONS section where the basic specifications for the research model are provided. This section includes the specification of requirements for qualified responses. The RESEARCH MODEL section describes the constructs involved and their associations. Finally, the RESEARCH HYPOTHESES section describes the hypotheses to be tested.

3.2 CONTEXT, SITUATION AND UNIT OF ANALYSIS RESEARCH SPECIFICATIONS

The context [Brannick, T. 2000] of this research is organizational settings operating with customer relationship processes based in a data warehouse. The type of situation [Brannick, T. 2000] or event that this research focuses on is customer relationship problem sense making. The unit of analysis in this research is the individual manager engaged in customer relationship processes (i.e. the context) in order to understand customer relationship problems (i.e. the situation). The profile of a qualified informant was defined in the following items (see the questionnaire in APPENDIX E: SURVEY QUESTIONNAIRE.v4):

- Actual dedication
  - The informant is engaged in their personal job in one of the following four CRM functions: Sales support, marketing support, customer service support, other CRM support (item #61).
  - The firm's data warehouse supports at least one of the following four CRM functions: Sales support, marketing support, customer service support, other CRM support (item #60).
  - The informant spends time enacting CRM problems (item #58 > 0).
  - The informant spends time searching/requesting customer relationship information in his/her firm's data warehouse (item #59 > 0). The items in the

behavioral section of the questionnaire (items #49 to item#54) are worded using the language "searching/requesting customer relationship information in their firm's data warehouse when you need to understand a customer relationship problem." In this way, I cover the following profiles:

- Individuals who need to understand a customer relationship problem and interface directly with the firm's data warehouse searching for information. This profile is labeled by the term 'researcher' (e.g. [Moorman and Austin 1995]). In the industry this profile is further segmented into 'power user', 'business user', and 'casual user' depending on the user's skills and the type of techniques he or she uses (e.g. [Gile 2002]).

- Individuals who need to understand a customer relationship problem and interface with a research agent who will interface directly with the firm's data warehouse in order to search for the requested information.

- Experience
  - The compound factor of (i) the informant's CRM experience (item #55), (ii) the informant's experience using the firm's data warehouse customer data (item #56), and (iii) time the firms' data warehouse supporting CRM (item #57) is at least 6 months (item #55 * item #56 * item #57 > 1).

3.3 RESEARCH MODEL

To the best of my knowledge, Bandura's SCT [1986] is the best general theory covering my research question *What are the inner and inter construct associations of the quality of data warehouse customer relationship data for problem enactment*? (see Table 3-1). This theory involves triadic associations among environmental latent constructs, cognitive latent constructs, and behavioral latent constructs.

- **Quality of data warehouse customer relationship data for problem enactment (QUALITY)**: Manager's perception on the extent to which the information derived from the available customer relationship data in the data warehouse fits with the manager's CRM cognitive needs for problem enactment (e.g. [Goodhue *et al.* 2000, Huang *et al.* 1998]).

SOCIAL COGNITIVE THEORY	CONSTRUCTS OF INTEREST IN RESEARCH QUESTION
Environmental constructs	• Quality of data warehouse customer relationship data for problem enactment (QUALITY)
Cognitive constructs	• Customer relationship problem sense making enhancements (ENACENH)
Behavioral constructs	• Customer relationship problem sense making information search mode orientation (MODE) • Customer relationship problem sense making information search heuristics orientation (HEUR)

Table 3-1 Mapping types of constructs of the SCT to the research model

. **Customer relationship problem sense making enhancements (ENACENH)**: Manager's self-belief about the extent to which the enactment of customer relationship problems has improved due to the data warehouse's customer relationship data (e.g. [Weick 2000, DeLone and McLean 1992, Weick 1993]).

Customer relationship problem sense making information search mode is the manager's information search behavior characterized by the type of inquiry. The information search modes considered in this research are intensity of data warehouse scanning search mode (SCAN) and intensity of data warehouse focused search mode (FOCUS) (e.g. [Huber 1991, Aguilar 1967]).

. **Customer relationship problem sense making information search mode orientation (MODE)**: Behavioral pattern in terms of the exhibited information search modes SCAN and FOCUS.

Customer relationship problem sense making information search heuristics is the manager's information search behavior characterized by a simplifying routine. The information search heuristics considered in this research are intensity of the data

warehouse availability search heuristic (AVBLE), intensity of the data warehouse representativeness search heuristic (REP), intensity of the data warehouse anchoring and adjustment search heuristic (AA), and intensity of the data warehouse positivy search heuristic (POSI) (e.g. [Hogarth 1987, Schwenk 1984, Hogarth and Makridakis 1981, Tversky and Kahneman 1974]).

- **Customer relationship problem sense making information search heuristics orientation (HEUR)**: Behavioral pattern in terms of the exhibited information search heuristics REP, AVBLE, AA and POSI.

The potential impact (e.g. mediation, group differences) of the following two variables is also part of this research:

- **CRM job function (JOBFUNC)**: CRM job function primarily performed by the informant. Four CRM job functions were studied: sales support, marketing support, customer services support and other CRM functions.

- **CRM data warehouse function (DWFUNC)**: CRM functions supported by the data warehouse. Four CRM job functions were studied: sales support, marketing support, customer services support and other CRM functions.

Following the collective contributions of SCT, and literature on data quality, CRM, information search behavior and problem solving, my research model is depicted in Figure 3.1. The constructs involved in the hypothesized internal structure of QUALITY and ENACENH are explained in the next section.

3.4 RESEARCH HYPOTHESES

3.4.1 Measurement Hypotheses

From my literature review, I have concluded that data quality is dimensional and contextual. More precisely, in the context of marketing activities, users' perceptions on information value are essentially interpreted in terms of instrumental utility. Collectivelly, this suggests that the latent construct quality of data warehouse customer relationship data for problem enactment (QUALITY) should have an

internal structure supporting this view (see Figure 3.2) by a data task utility dimension (TASKUT) that provides instrumental utility for the context (i.e. CRM).

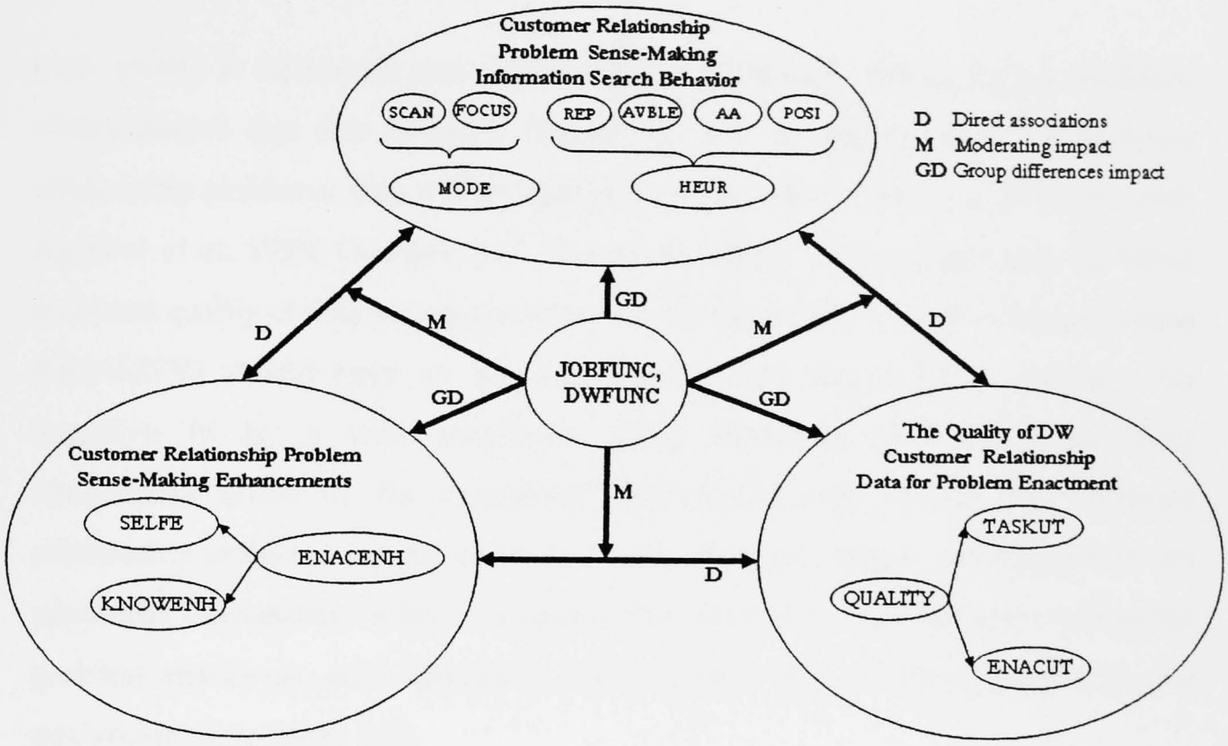


Figure 3.1 Detail-Level Research Model

Furthermore, in marketing information processing literature, data insights refers to the attribution of cognitive utility to data. Customer insights and competitor insights are the two main types of data insights in the context of this research. This suggests that the data task utility dimension (TASKUT) should have an internal structure supporting this view by a customer insights dimension (CUST) and a competitor insights dimension (COMP).

- **Data task utility (TASKUT)**: Manager's perception on the extent to which the data warehouse customer relationship data has instrumental utility for the task at hand. In our case, the tasks are in the context of CRM.

- **Customer insights (CUST)**: Manager's perception on the extent to which the data warehouse customer relationship data is a source of customer data insights (e.g. [Day 1994]).

- **Competitor insights (COMP)**: Manager's perception on the extent to which the data warehouse customer relationship data is a source of competitor data insights (e.g. [Day 1994]).

Data quality is contextual and situational (e.g. [Brannick 2000]). In the situation-within-context that this research focuses on, that is, making sense of customer relationship problems, data quality should have cognitive utility (e.g. [Vessey 1991, Agarwal *et al.* 1996, Goodhue and Thompson 1995]). This suggests that the latent construct quality of data warehouse customer relationship data for problem enactment (QUALITY) should have an internal structure (see Figure 3.2) supporting this cognitive fit by a data enactment utility dimension (ENACUT) providing instrumental utility for the considered situation-within-the context (i.e. customer relationship problem enactment). Furthermore, previous research has found that this advocated instrumental utility is supported by dimensions like customer relationship problem enactment, data integration, data richness, data trustworthiness, and data equivocality (see Table 3-2).

Overall, this suggests that the data enactment utility dimension (ENACUT) should have an internal structure supporting this view by a customer relationship problem enactment dimension (ENAC), data integration dimension (INT), data richness dimension (RICH), data trustworthiness dimension (TRUST), and data equivocality dimension (EQUI).

- **Data enactment utility (ENACUT)**: Manager's perception on the extent to which the data warehouse customer relationship data has enactment utility for the specific confronted situation. In our case, the situation is enacting customer relationship problems.

- **Customer relationship problem enactment (ENAC)**: Manager's perception on the extent to which the data warehouse is a source of applicable and helpful customer relationship data for enacting customer relationship problems (e.g. [Fedorowicz and Lee 1998, Bailey and Pearson 1983, Venkatesh and Davis 2000]).

DATA ENACTMENT UTILITY (ENACUT) DIMENSIONS	REFERENCES
Customer relationship problem enactment (ENAC)	(e.g. [Fedorowicz and Lee 1998, Bailey and Pearson 1983, Venkatesh and Davis 2000, Moenaert and Souder 1996, Jenkins and Ricketts 1985, Wang and Strong 1996, O'Reilly III 1982, Huang *et al.* 1998, Swanson 1987, Haley 1997, Streufert 1973, Boynton 1987, King and Epstein 1983]),
Data integration (INT)	(e.g. [Lawrence and Lorsch 1967, Goodhue *et al.* 2000, Orasanu and Connolly 1995, March and Olsen 1976, Weick 1976, Boynton and Zmud 1987, Wetherbe 1991])
Data richness (RICH)	(e.g. [Weick 2000, Moorman 1995, Daft and Lengel 1984])
Data trustworthiness (TRUST)	(e.g. [Grooms 2001, Jarke *et al.* 1999, Moenaert and Souder 1996, Moorman *et al.* 1992, Wixom and Watson 2001, Song *et al.* 2000, Swanson 1987, Srinivasan 1985, Seddon and Kiew 1994, O'Reilly III 1982, Kettinger and Lee 1994, Venkatesh and Davis 2000, Choo 1993])
Data equivocality (EQUI)	(e.g. [Goodhue *et al.* 2000, Jarke *et al.* 1999, Franz 1999, Brannick 1998, Swanson 1987, Daft and Lengel 1986, Daft and Weick 1984, Daft and Macintosh 1981, Weick 1979])

Table 3-2 Literature addressing data enactment utility

- **Data integration (INT)**: Manager's perception on the extent to which the data warehouse customer relationship data is normalized in terms of data definitions and logical data structures (e.g. [Goodhue *et al.* 1992, Goodhue *et al.* 2000, Peltier *et al.* 1998, Codd 1972]).

- **Data richness (RICH)**: Manager's perception on the extent to which the data warehouse customer relationship data is a source of customer relationship information (e.g. [Daft and Lengel 1984]).

- **Data equivocality (EQUI)**: Manager's perception on the extent to which the data warehouse customer relationship data favors more than one interpretation for the enactment of customer relationship problems (e.g. [Goodhue *et al.* 2000, Jarke *et al.* 1999, Weick 1979, Daft and Lengel 1984, Daft and Lengel 1986, Swanson 1987]).

- **Data trustworthiness (TRUST)**: Manager's perception on the extent to which the data warehouse customer relationship data is regarded as true and credible evidence for the enactment of customer relationship problems (e.g. [Jarke *et al.* 1999]).

Hypothesis 1: The quality of data warehouse customer relationship data for problem enactment QUALITY is a high order latent construct with two dimensions including (i) a context dependent (i.e. CRM) data task utility dimension TASKUT providing instrumental utility for the context's tasks, and (ii) a situation-within-context dependent (i.e. customer relationship problem enactment) data enactment utility dimension ENACUT providing instrumental utility for the considered situation-within-context.

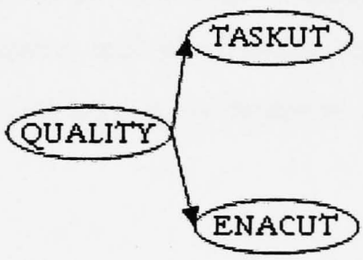


Figure 3.2 Disagregated research model for QUALITY

The theory of sense making posits that (i) reality is an ongoing accomplishment (e.g. it emerges from efforts to create order and make retrospective sense of what occurs) [Weick 1993], and that (ii) to make sense of an event (e.g. a customer relationship problem) means that heedful interrelating connects sufficient individual knowledge with situational demands [Weick 1993]. A problem statement is the result of a sense making process known as enactment (e.g. [Weick 2000]). Therefore, enacting customer relationship problems is one of the key processes in making sense of customer relationship problems. Because data quality is the cornerstone of this

research, I am interested in understanding whether or not problem sense making enhancements can be attributable to data quality. According to the SCT, proficiency in performing a task will be heavily influenced by self-efficacy beliefs and enactive mastery experience (i.e. personal attainments), which is the most influential source of efficacy information [Bandura 1997]. Performance attainments have an impact on self-efficacy beliefs [Bandura 1997]. Furthermore, an individual's cognitive appraisal and integration of these experiences ultimately determine self-efficacy [Bandura 1982].

This suggests that the latent construct customer relationship problem sense making enhancements (ENACENH) should have an internal structure (see Figure 3.3) supporting the links between (i) an enactment mastery experience knowledge enhancement dimension (KNOWENH) reporting self-beliefs on personal attainments in the considered situation-within context (i.e. customer relationship problem enactment), and (ii) an enactment mastery expectancy self-efficacy dimension (SELFE) reporting the self-efficacy beliefs in prospective situations-within context (i.e. customer relationship problem enactment).

. **Customer relationship problem enactment knowledge enhancement (KNOWENH)**: Manager's self-belief about the extent that one's skills enacting customer relationship problems have improved due to the data warehouse customer relationship data.

. **Customer relationship problem enacting self-efficacy (SELFE)**: The strength in the manager's self-belief in one's capabilities to execute given types of performances enacting prospective customer relationship problematic situations (e.g. [Bandura 1997, 1986]).

Hypothesis 2: Customer relationship problem sense making enhancements ENACENH is a high order latent construct with two dimensions including the following two situation-within-context dependent (i.e. customer relationship problem enactment) dimensions: (i) an enactive mastery experience knowledge enhancement

dimension KNOWENH reporting self-beliefs on knowledge enhancements in the considered situation-within-context, and (ii) an enactive mastery expectancy self-efficacy dimension SELFE reporting self-efficacy beliefs in prospective situations-within-context.

Previous research supports the notion that cognitive structures holding knowledge can be characterized by their integrative complexity (e.g. [Sullivan and Weaver 2000, Wang and Chan 1995, Feist 1994, McFadzean 1996, Stone 1994]). Anderson's [1993] taxonomy of knowledge and subsequent research (e.g. [Porter and Inks 2000, Leidner, Carlsson, Elam and Corrales 1999, Campbell 1994]) support a declarative knowledge and a procedural knowledge. Overall, this suggests that the advocated customer relationship problem enactment knowledge enhancement dimension (KNOWENH) should have an internal structure supporting these findings by a customer relationship problem declarative knowledge enhancement dimension (DKNOWE), a customer relationship problem enactment procedural knowledge enhancement dimension (PKNOWE), and a customer relationship problem integrative complexity enhancement dimension (ICPLXE).

- **Customer relationship problem declarative knowledge enhancement (DKNOWE)**: Manager's self-belief about the extent that knowledge on customer relationship problems has improved due to the data warehouse customer relationship data. Declarative knowledge is the set of stored situational cues and facts (e.g. types of customers and selling situations) which allows the manager to recognize and categorize customer relationship situations (e.g. [Porter and Inks 2000, Leidner *et al.* 1999, Campbell 1994]).

- **Customer relationship problem enactment procedural knowledge enhancement (PKNOWE)**: Manager's self-belief about the extent that one's skills enacting customer relationship problems have improved due to the data warehouse customer relationship data. Procedural knowledge consists of routines, actions, strategies, or heuristics that apply to a task domain (e.g. [Porter and Inks 2000, Campbell 1994]).

. **Customer relationship problem integrative complexity enhancement (ICPLXE)**: Manager's self-belief about the extent that the integrative complexity of the cognitive structures about customer relationship problems has improved due to data warehouse customer relationship data. Integrative complexity is the level of comprehensiveness -i.e. number of factors in the cognitive structure- and connectedness -i.e. links among the factors in the cognitive structure (e.g. [McFadzean 1996, Wang and Chan 1995, Stone 1994]).

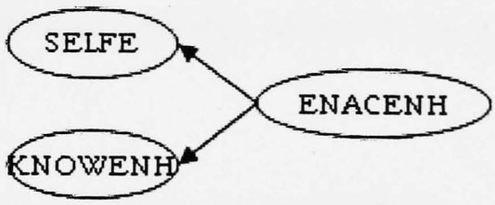


Figure 3.3 Disagregated research model for ENACENH

Previous empirical research has found differences in terms of people's behavioral patterns in their information search modes (e.g. [Gonzalez 2001, Maier, Rainer Jr. and Snyder 1997, Vandenbosch and Higgins 1996, Choo 1993, Boynton 1987]) and in the information search heuristics (e.g. [Lau and Redlawsk 2001, Evans 1989, Folkes 1988, Tversky and Kahneman 1974]). This means that I should not expect to find in this research that (i) focus and scan are evenly used and (ii) that the four selected heuristics are evenly used either. Therefore, this suggests that we should expect a number of different behavioral patterns in terms of (i) the proportion of scanning versus focus search, and (ii) the proportion in simplifying routines. However, the literature does not suggest any a-priori expectation on the predominant information search mode or information search heuristic.

. **Intensity of data warehouse scanning search mode (SCAN)**: Manager's perceptions on the personal's amount of effort in scanning the data warehouse making sense of customer relationship problems. Scanning search is the proactive and exploratory information search behavior mode people exhibit when they browse through information without a particular problem to solve (e.g. [Maier *et al.* 1997, Boynton 1987]).

- **Intensity of data warehouse focused search mode (FOCUS)**: Manager's perceptions on the personal's amount of effort in focused search on the data warehouse making sense of customer relationship problems. Focused search is the reactive and directed information search behavior mode people exhibit when they are looking for information specific to a problem to be addressed or question to be answered (e.g. [Maier *et al.* 1997, Boynton 1987]).

Hypothesis 3.1: There are different behavioral patterns MODE in terms of the exhibited customer relationship problem sense making information search modes SCAN and FOCUS.

- **Intensity of the data warehouse representativeness search heuristic (REP)**: Manager's perceptions on the personal's amount of effort using representativeness search heuristic in the data warehouse in order to make sense of customer relationship problems. Representativenes heuristic search refers (i) To assessing the probability of a situation as representative of a category (e.g. a 'price-lowering by a competitor' situation can have common information with an 'attempt action to gain market-share' pattern), or (ii) to make generalizations based on new information about a sample - i.e. the sample is representative of a large population- (e.g. to assess the national market success of a new product line based on the data likehood ratio of a test market). This heuristic implies that the search ends when there is a satisficing fit between information about a situation and information about a category (e.g. [Chi and Fan 1997, Wright 1980]).

- **Intensity of the data warehouse availability search heuristic (AVBLE)**: Manager's perceptions on the personal's amount of effort using availability search heuristic in the data warehouse in order to make sense of customer relationship problems. Availability heuristic search refers to assessing the probability of a situation as a function of prior situations (e.g. a marketer considering a series of occurrences of actual costs incurred in past editions of a campaign when estimating the cost for a new edition of such campaign). This heuristic implies the search for easily accessible information about relevant precedents. The search ends once recent,

salient information about a relevant precedent is found (e.g. [Chi and Fan 1997, Wright 1980]).

- **Intensity of the data warehouse anchoring and adjustment search heuristic (AA)**: Manager's perceptions on the personal's amount of effort using anchoring and adjustment search heuristic in the data warehouse in order to make sense of customer relationship problems. Anchoring and adjustment heuristic search refers essentially to the trial and error method (e.g. a marketer setting the price of a product starting with a baseline price and making a number of impact analyses in several of the cost components). This heuristic implies a recursive process and each step involves a search for additional information and an adjustment of the previous assessment. The search ends once the adjustments are not improving an implicit/explicit value function on the information found (e.g. [Chi and Fan 1997]).

- **Intensity of the data warehouse positivy search heuristic (POSI)**: Manager's perceptions on the personal's amount of effort using positivy search heuristic in the data warehouse in order to make sense of customer relationship problems. Positivy heuristic search refers to confirming the probability of a situation using the trial and error method (e.g. a marketer looking for issues in accounts that were predicted to have a high risk). This heuristic implies a search for information that is fundamentally consistent with existing beliefs, theories and cognition. The search ends once the information found confirms the probability of a situation (e.g. [Evans 1989]).

Hypothesis 3.2: There are different behavioral patterns HEUR in terms of the exhibited customer relationship problem sense making information search heuristics REP, AVBLE, AA and POSI.

3.4.2 Model Structure Hypotheses

SCT posits a triadic causation model in which behavior, cognitions, and the environment all influence each other in a dynamic fashion [Bandura 1997]. This suggests that there is a triadic association among (i) the quality of data warehouse customer relationship data for problem enactment, (ii) customer relationship problem

sense making enhancements, and (iii) customer relationship problem sense making information search behavior.

Vessey's theory of cognitive fit (e.g. [1991]), Simon's human information processing theory (e.g. [1976]), and attribution theory (e.g. [Kelley 1967]) collectively suggest that there is causation between (i) the quality of data warehouse customer relationship data for problem enactment constructs, and (ii) customer relationship problem sense making enhancements constructs (see Figure 3.4).

Hypothesis 4.1: There is an association between the quality of data warehouse customer relationship data for problem enactment QUALITY, and customer relationship problem sense making enhancements ENACENH.

Bandura's SCT [1986], Simon's human processing information theory (e.g. [Simon 1976]) and Kolb's learning theory (e.g. [Kolb 1974]) collectively suggest that there is causation between (i) customer relationship problem sense making enhancements constructs, and (ii) customer relationship problem sense making information search behavioral constructs (see Figure 3.4).

Hypothesis 4.2: There is an association between customer relationship problem sense making information search behavioral mode patterns MODE and customer relationship problem sense making enhancements ENACENH.

Hypothesis 4.3: There is an association between customer relationship problem sense making information search behavioral heuristic patterns HEUR and customer relationship problem sense making enhancements ENACENH.

Expectancy and adaptive behavior theories (e.g. [Fishbein 1967, Fishbein and Ajzen 1975, and Payne 1976]), Sinkula's organizational learning theory on marketing tasks [Sinkula 1994], and attribution theory (e.g. [Kelley 1967]) collectively suggest that there is causation between (i) the quality of data warehouse customer relationship

data for problem enactment constructs, and (ii) customer relationship problem sense making information search behavioral constructs (see Figure 3.4).

Hypothesis 4.4: There is an association between customer relationship problem sense making information search behavioral mode patterns MODE and the quality of data warehouse customer relationship data for problem enactment QUALITY.

Hypothesis 4.5: There is an association between customer relationship problem sense making information search behavioral heuristic patterns HEUR and the quality of data warehouse customer relationship data for problem enactment QUALITY.

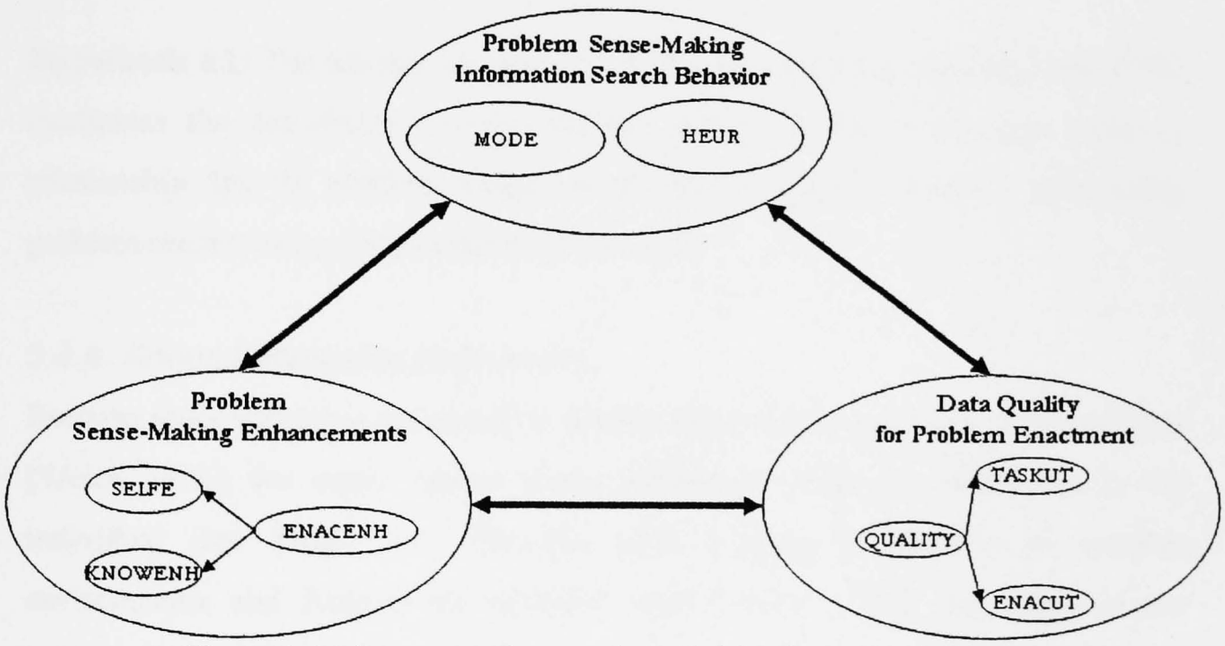


Figure 3.4 Disagregated research model for direct associations

3.4.3 Moderating Hypotheses

Customer relationships management tasks, as part of the marketing function, are traditionally grouped into jobs [Kotler 1988]. Therefore, I am interested in understanding if there are differences in the association between the quality of data warehouse customer relationship data for problem enactment and customer relationship problem sense making enhancements due to different environmental situations like CRM job function (see Figure 3.1).

Hypothesis 5.1: CRM job function JOBFUNC moderates the association between (i) the quality of data warehouse customer relationship data for problem enactment

QUALITY, and (ii) customer relationship problem sense making enhancements ENACENH.

To the best of my knowledge, there is no serious empirical research that has studied differences due to the number of CRM functions supported by the data warehouse in the constructs of this research model. Therefore, I am interested in understanding if there are differences in the association between the quality of data warehouse customer relationship data for problem enactment and customer relationship problem sense making enhancements due to the number of CRM functions supported by the data warehouse (see Figure 3.1).

Hypothesis 6.1: The number of suported CRM data warehouse functions DWFUNC moderates the association between (i) the quality of data warehouse customer relationship data for problem enactment QUALITY, and (ii) customer relationship problem sense making enhancements ENACENH.

3.4.4 Group Differences Hypotheses

Because sense making is grounded in identity construction and based on plausibility [Weick 1995], we might expect group differences owing to differences in the individual (see Figure 3.1). Because sense making is enactive of sensible environments and focused on extracted cues [Weick 1995], we might expect individuals in different jobs and/or environments to make sense of things differently (e.g. [Taylor 1999]).

Hypothesis 5.2: CRM job function JOBFUNC produces group differences in the quality of data warehouse customer relationship data for problem enactment QUALITY.

Hypothesis 5.3: CRM job function JOBFUNC produces group differences in customer relationship problem sense making enhancements ENACENH.

Likewise, we should expect individuals in different jobs to have different information search behaviors (e.g. [Vandenbosch and Huff 1997]). Particularly, it is expected to

find that individuals performing the marketing support job function exhibit higher levels of intensity of data warehouse scanning search mode (SCAN) than individuals performing other CRM functions because marketing support is typically considered a more strategic job than sales support and customer support (e.g. [Kotler 1988]).

Hypothesis 5.4: CRM job function JOBFUNC produces group differences in customer relationship problem sense making information search modes SCAN and FOCUS.

Hypothesis 5.5: CRM job function JOBFUNC produces group differences in customer relationship problem sense making information search heuristics REP, AVBLE, AA, and POSI.

In the industry, supporting more than one CRM function is associated with high data quality by enterprise data warehouse merchants. There is an optimistic assumption that (i) the quality of data warehouse customer relationship data for problem enactment grows with the number of supported CRM functions, (ii) data warehouses supporting more CRM functions enable more enhancements to the users than data warehouses supporting less CRM functions, and (iii) data warehouses supporting more CRM functions enable more scanning than data warehouses supporting less CRM functions.

Hypothesis 6.2: The number of supported CRM data warehouse functions DWFUNC produces group differences in the quality of data warehouse customer relationship data for problem enactment QUALITY.

Hypothesis 6.3: The number of supported CRM data warehouse functions DWFUNC produces group differences in customer relationship problem sense making enhancements ENACENH.

Hypothesis 6.4: The number of supported CRM data warehouse functions DWFUNC produces group differences in customer relationship problem sense making information search modes SCAN and FOCUS.

Hypothesis 6.5: The number of supported CRM data warehouse functions DWFUNC produces group differences in customer relationship problem sense making information search heuristics REP, AVBLE, AA, and POSI.

3.5 SUMMARY OF CONCEPTUALIZATION CHAPTER

This chapter has provided a comprehensive description of the hypotheses derived from the research model. The sections in this chapter have followed the structure of the research model. The research model was disaggregated into several research submodels in order to facilitate the understanding of the different aspects involved.

Hypotheses cover a comprenhesive test of the internal structure of involved constructs, their associations, and impact of moderating variables.

4. RESEARCH METHODOLOGY

4.1 INTRODUCTION

Research design is the term that the literature uses for the plan that the researcher proposes to follow when conducting the research (e.g. [Hair Jr. *et al.* 2003, Remenyi, Williams, Money and Swartz 1998]). Therefore, research design should cover the entire life cycle of such research. Research strategy is the basic philosophical orientation of the research [Remenyi *et al.* 1998]. Research analysis is the process that handles the verification or falsification of the research hypotheses (e.g. [Hair Jr. *et al.* 2003]). The adopted research analysis must be consistent with the selected research strategy. From my literature review I have concluded that the terms analytical phase [Hair Jr. *et al.* 2003], research method [Easterby-Smith, Thorpe and Lowe 1991], and research tactic [Remenyi *et al.* 1998] all refer to the same concept. I selected the terms research strategy, research design, and sample data analysis for use in this paper.

This chapter starts with the RESEARCH STRATEGY section describes the high level research decisions while the RESEARCH DESIGN section covers the details of the operationalization. I put special care in presenting the sections in an order that would facilitate understanding of the research process. For example, it was very difficult to find a way of explaining the different versions of the survey. I addressed this challenge by (i) documenting the first and fourth (last) versions in a section and creating an appendix with details of all the steps and (ii) qualifying each item with a suffix indicating the operationalization version to which it belongs.

4.2 RESEARCH STRATEGY

This research is both of an exploratory and confirmatory nature (e.g. [Pinsonneault and Kraemer 1993, McGrath 1979]). It is exploratory since it involves (i) known high order latent constructs with little knowledge of their internal structures, and (ii) the adaptation of existing scales in order to measure such latent constructs. It is confirmatory since it involves theoretically known relationships. Rather than a strict dichotomy of exploratory vs. confirmatory, this research can be thought of as an ordered progression [Anderson and Gerbing 1988].

As explained in McGrath [1982], a research strategy choice will meet only one of the following three desiderata [McGrath 1982]: Generalizability of research findings with respect to populations; precision with regard to measurement, manipulation and control; and existential realism. Because the research model has variables of a different nature (i.e. cognitive, behavioral and environmental), I decided the following:

- Generalizability should be my first priority. From the repertoire of positivistic options for making observations, I selected the large-scale survey method. Using a questionnaire has become an increasingly common way of investigating managerial behavior, organizational properties an individual believes as a way of overcoming lack of generalizability of single cases, and the lack of a real-life feel of laboratory experiments (e.g. with students as surrogates of managers).
- I should partially mitigate the weaknesses in precision and realism with the following research actions: (i) A world-wide survey in order to increase realism of the context variables, and (ii) focus groups at different stages of the research process in order to increase realism on the environmental and behavioral variables (see Table 4-1) in order to provide a qualitative perspective to the validations to be performed. This approach was presented in [Abril 2002]. I considered case studies as an alternative interpretative method in conjunction with a large-scale survey (i.e. as opposed to focus groups) and concluded that this type of method would provide a less significant contribution than focus groups. I presented my findings in this respect in [Abril 2002].

Analysis of Strenghts and Weaknessess of Research Method Choices

	Survey Expectations			Focus Group Expectations		
Type	Generalizability	Precision	Realism	Generalizability	Precision	Realism
Behavioral	H		M			M-H
Cognitive	H		M			
Environmental	H		M			H
Meet desiderata	Caution	Weakness				

Contribution of the Focus Group to Realism

Table 4-1 Selected data collection methods

- I decided to manage this research as a project. The Project Management Body of Knowledge [Project Management Institute 2000], PMBOK in short, developed and issued by the Project Management Institute, describes the knowledge of proven practices which are widely applied as well as knowledge of innovative practices which have seen less use. I confirmed the viability of the PMBOK for academic research in a paper that I presented at the 2002 International Symposium on Research Methods, International Symposium on Research Methods (Las Vegas, N) [Abril 2002]. The fundamental consideration supporting the application of the PMBOK in research, my argument in such paper, is that research is not routine and is instead characterized by the following distinctive attributes: unique, temporary, performed by people with limited resources, and with a criteria for defining success.
- I was the overall owner of the project and sole research agent conducting the research. Therefore, I am the only person responsible for what is contained in this research.
- My supervisors acted as approvers. This meant that I proceeded on a course of action always under the explicit approval of all my supervisors. In simple terms, all research decisions were agreed to by my supervisors.
- I reduced complexity in this research project by defining its life cycle as a number of subprojects. My criteria for identifying research subprojects follows Remenyi *et al.* guidelines [1998]: Literature review, assessment of established theoretical frameworks, assessment of grounded theory in case of weak theoretical basis, theoretical conjecture and hypotheses formulation, operationalization, sampling, testing and analysis, confirmation of theory and development of further/refined theory. Additional insights for the data collection plan came from the market research literature (e.g. [Churchill Jr. 1991]). Overall, the life cycle follows a nonlinear incremental approach and requires frequent interactions. The life cycle of this research is exhibited in Figure 4.1.

According to the PMBOK, every project has three management drivers that will determine its success or failure: quality, time, and cost [Project Management Institute 2000]. This means that success requires closing the project meeting quality research requirements, on time and on budget. Reality also requires you to select one driver

and accommodate the other two accordingly. My choice was quality at the expense of time and cost.

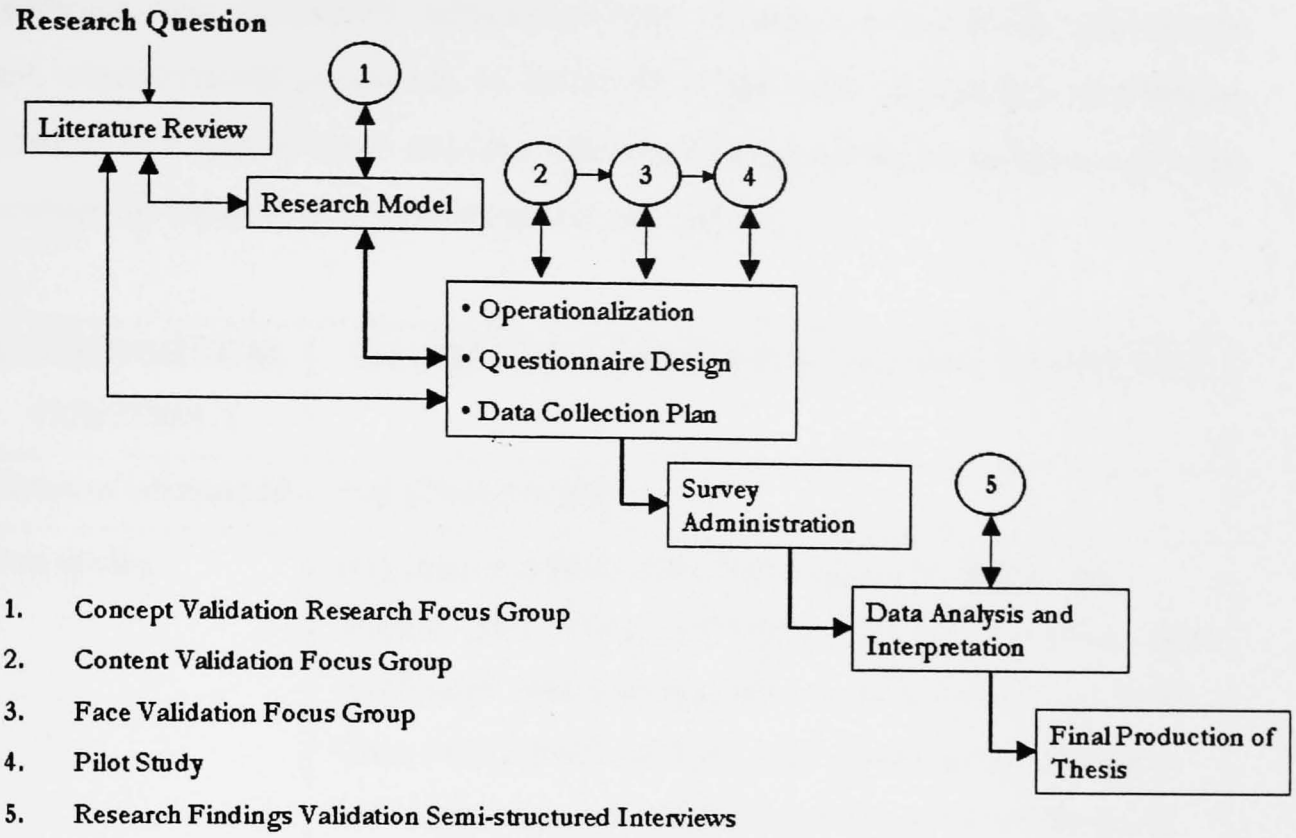


Figure 4.1 Life cycle of this research

Finally, PMBOK [Project Management Institute 2000] guidelines in the area of communications were systematically applied:

- Status reporting was formal and took the form of e-mails reporting status and progress to my supervisors.
- Intermediate deliverables were discussed in direct conversations with scheduled conference calls. The format used was PowerPoint slides were I inserted the relevant pieces of data (e.g. SAS, SPSS and AMOS tables).

4.3 RESEARCH DESIGN

4.3.1 Subjective Judgments as Valid Evidence Collection Procedure

The literature on information management stresses the crucial importance of the opinions of the user of the service over the opinions of the supplier of such service (e.g. [Deshpande, Farley and Webster Jr. 1993]). Past research based on subjective judgment is endless. Some environmental constructs studied using subjective

judgments are: marketing practitioners' perceptions and understanding of the problematic nature of information, data quality, information credibility, data accesibility, data availability, information load, information usefulness, information value, organizational ownership of information, and user information satisfaction. Therefore, it is reasonable to consider subjective judgments as an evidence collection procedure for the quality of data construct (see Table 4-2).

ENVIRONMENTAL CONSTRUCT	REFERENCES SUPPORTING SUBJECTIVE JUDGMENTS
Nature of information	(e.g. [Brannick 2000])
Data quality	(e.g. [Low and Mohr 2001, Van Birgelen, de Ruyter and Wetzels 2001, Wixom and Watson 2001, Chen, Soliman, Mao and Frolick 2000, Kettinger and Lee 1994, Huang *et al.* 1998, Choo 1993, Jobber and Watts 1988, Boynton 1987, Swanson 1987, Srinivasan 1985, Bailey and Pearson 1983, O'Reilly III 1982, Zmud 1978, Lucas Jr. 1973])
Information credibility	(e. g. [Wanzenried and Powell 1993])
Data accessibility	(e. g. [Choo 1993, Boynton 1987, O'Reilly III 1982])
Data availability	(e. g. [Liang 1986, Leidner *et al.* 1999])
Information load	(e. g. [Goslar, Green and Hughes 1986, Chorba and New 1980, Lucas Jr. and Nielson 1980, Schroeder and Benbasat 1975])
Information usefulness	(e. g. [Mahmood and Medewitz 1985, Lucas Jr. 1981])

Table 4-2 Literature supporting subjective judgements on environmental constructs

In relation to the measurement of managerial cognitive characteristics, Albert Bandura's guidelines [Bandura 1997, Pajares, Hartley and Valiante 2001] advocate for subjective judgment measuring self-beliefs. Some cognitive constructs studied using subjective judgments are problem-solving cognitive style, problem-solving self-appraisal, ability to identify strategic opportunities or problems, problem-solving confidence, problem-solving self-efficacy, sales self-efficacy, and computer self-

efficacy. Therefore, it is reasonable to consider subjective judgments as an evidence collection procedure for the knowledge enhancements constructs (see Table 4-3).

COGNITIVE CONSTRUCT	REFERENCES SUPPORTING SUBJECTIVE JUDGMENTS
Problem-solving cognitive style	(e. g. [Averett 1991, Jabri 1991])
Problem-solving self-appraisal	(e. g. [Neal and Heppner 1986])
Ability to identify strategic opportunities or problems	(e. g. [Goul, Shane and Tonge 1986])
Problem-solving confidence	(e. g. [Heppner, Cook, Wright and Johnson Jr. 1995, Heppner 1988])
Problem-solving self-efficacy	(e. g. [Heppner, Cooper, Mulholland and Wei 2001, Wolf 1997])
Sales self-efficacy	(e. g. [Srivastava, Strutton and Pelton 2001, Silver 2000])
Computer self-efficacy	(e. g. [Compeau, Higgins and Huff 1999, Munro, Huff, Marcolin and Compeau 1997])

Table 4-3 Literature supporting subjective judgements on cognitive constructs

In relation to the measure of information search behavior, subjective judgment has been a common evidence collection procedure of choice although there are other alternatives [Easterby-Smith *et al.* 1991]. Some examples of studies using subjective judgments are the following: information generation, information acquisition/search/scanning, information implementation/usage, information dissemination, market orientation, problem-solving strategies, problem finding, and problem formulation. Therefore, it is reasonable to consider subjective judgments as an evidence collection procedure for the information search behavior constructs (see Table 4-4).

BEHAVIORAL CONSTRUCT	REFERENCES SUPPORTING SUBJECTIVE JUDGMENTS
Information generation	(e. g. [Sinkula, Baker Jr.and Noordewier 1997])
Information acquisition/search/ scanning	(e. g. [Gonzalez 2001, Akgun 2001, Moorman 1995, Averett 1991, Ostroff and Kozlowski 1992, Burky 1990, Sinkula and Hampton 1988, Boynton 1987])
Information implementation/usage	(e. g. [Van Birgelen *et al.* 2001, Akgun 2001, Moorman 1995, Choo 1993, Moorman *et al.* 1992, Swanson 1987, Deshpande and Zaltman 1984 and 1982, Benbasat, Dexter and Masulis 1981, Schroeder and Benbasat 1975])
Information dissemination	(e. g. [Akgun 2001, Sinkula *et al.* 1997, Moorman 1995])
Market orientation	(e. g. [Jaworski and Kohli 1993])
Problem-solving strategies	(e. g. [Heppner *et al.* 2001])
Problem finding	(e. g. [Sayeed and Brightman 1994])
Problem formulation	(e. g. [Chapman 1989, Cowan 1988])

Table 4-4 Literature supporting subjective judgements on behavioral constructs

4.3.2 Scale Validity Plan

By scale validity x of a construct X, it is meant that the scale x is a valid measure of the construct X (i.e. has X been successfully operationalized by x?) [Cronbach and Meehl 1955, Easterby-Smith *et al.* 1991]. This can be assessed by ensuring the following steps [Straub 1989]:

. The definition of the construct X has been done properly [Schwab 1999]. The definitions of the QUALITY constructs in my research model (see Figure 3.1) were addressed in a concept validation research focus (see Concept Validation Research Focus section) and in a subsequent conference paper [Abril 2001]. The other constructs in my research model have been well studied in the literature and did not require further concept validation.

. Scale x conforms to the conceptual definition of X by avoiding systematic errors (i.e. deficiencies and contamination [Schwab 1999]). This requires demonstration that

the scale x is consistent with the definition of the construct X [Hair Jr., Anderson, Tatham and Black 1998, Schwab 1999]. This can be achieved with content validity, face validity and measurement administration.

- Content validity x-X: This entails deciding whether the scale x covers all the dimensionalities of the construct X [Reaves 1992] (i.e. there are no deficiencies) and that it is not contaminated with capturing characteristics not specifically included in the definition of X [Schwab 1999]. According to [Reaves 1992, Schwab 1999], people who are experts on the construct X under scrutiny should be asked whether x seems to have content validity to them. Content validity for all the scales in the research model was addressed in a content validation research focus (see Content Validation Focus Group).
- Face validity x-X: This entails deciding whether the scale x looks valid as a measure of the construct X on the face of it. According to [Reaves 1992, Schwab 1999] people who are not particularly expert on the construct X under scrutiny should be asked about whether x seems to have face validity to them. Face validity for all the scales in the research model was addressed in a face validation research focus (see Face Validation Focus Group).
- Anything that influences scores can influence validity. The administrative environment (e.g. confidentiality [ASA 1999]) of the measure may also influence errors. See Survey Administration section for details about the administrative actions that were taken in order to avoid influencing the scores.

- Unidimensionality. The items in a summated scale should be unidimensional, meaning that they are strongly associated with each other and represent a single construct [Hair Jr. *et al.* 1998, Bagozzi, Yi and Phillips 1991]. The test of unidimensionality is that each item should load highly on a single factor. Unidimensionality of QUALITY and ENACENH was addressed when applying factor analysis and exploratory factor analysis.

- Demonstrate that the observations obtained conform to the conceptual definition of the construct X using the following validation procedures [Schwab 1999, Hair Jr. *et al.* 1998]:
  - Convergent validity x-X (also known as criterion validity, predictive validity [Schutt 1996] and concurrent validity [Reaves 1992]): Confirmation of the

validity of the scale x as a measure of the construct X by comparing the scores of x with the results yielded at the same time (i.e. concurrently) by another scale y about the construct Y [Easterby-Smith *et al.* 1991] and finding a high correlation [Hair Jr. *et al.* 1998]. Evidence of convergent validity adds to a researcher's confidence in the construct validity of measures [Schwab 1999]. Convergent validity of QUALITY and ENACENH was addressed when applying correlation analysis, exploratory factor analysis.

- Discriminant validity x-X y-Y [Schwab 1999]): It is confirmed when scores from measures x and y do not converge (i.e. low correlation [Hair Jr. *et al.* 1998]). In other words, it provides information about whether scores from x are unique with respect to X rather than if they are contaminated by another construct Y (i.e. that x is unique to X and therefore not measuring also a second construct). Discriminant validity is particularly important when developing a measure x of a new construct X that may be redundant with another construct Y which has been more thoroughly researched. Therefore, a measure x of proposed construct X should show evidence of discriminant validity with measures of existing constructs. Discriminant validity of QUALITY and ENACENH was addressed when applying exploratory factor analysis.

4.3.3 Concept Validation Research Focus

Initially, my research question only included the latent construct quality of data warehouse customer relationship data for problem enactment (QUALITY) and perceptions on the value provided by the data warehouse. I wanted (i) to validate my perception of the importance of my research question from a management perspective, (ii) to validate the conceptual clarity of the definition "Manager's perception on the extent to which the information derived from the available customer relationship data in the data warehouse fits with the manager's CRM cognitive needs for problem enactment" of the latent construct quality of data warehouse customer relationship data for problem enactment (QUALITY), and (iii) get some clues about the potential constructs that would be relevant in my research model. Therefore, I conducted an exploratory field intervention by arranging a focus group following the recommended guidelines of Stewart and Shamdasani [1998] of planning, selecting participants, moderating, data collection and principles of analysis. This study was

conducted on September 26, 2000 during the annual international conference that the user community of NCR's data warehouses organizes to address data warehouse and CRM subjects. Results of this exploratory research (see APPENDIX A: CONCEPT VALIDATION RESEARCH FOCUS GROUP for details) were coded with an identifier #1 to #9.

Result #3 (i.e. organizational financial performance measures were recommended although such measures were not available) suggets an inconsistency between desired status and reality. I conclude that participants experience difficulties implementing organizational performance measures for evaluating the contribution of their data warehouses and that this should be done through some subrogates. Results #2 and #4 (i.e. data quality is a recommended measure for the value contribution of a data warehouse), and #5 (i.e. QUALITY definition was clear) reveal a consistent interest by the participants in data quality as a factor which impacts the contribution of their data warehouses. Therefore, I found that practitioners support the inclusion of the latent construct quality of data warehouse customer relationship data for problem enactment (QUALITY) in my research question. Results #6 (i.e. low skills in information management is a barrier to promoting usage of the information derived from a data warehouse) and #9 (i.e. types of queries) revealed a consistent interest by the participants in information search behavior which was not part of my initial research question. Result #7 (i.e. knowledge deficiencies about the business process is a barrier to promoting usage of the information derived from a data warehouse) indicated the relevance for including cognitive benefits as part of my research question. Result #8 (i.e. job function influences the usage of the information derived from a data warehouse) indicated the convenience for considering moderator variables in this research. In general, the results from my exploratory research supported the managerial significance of my research question. The main results of this exploratory research were:

- I confirmed the conceptual clarity of the latent construct quality of data warehouse customer relationship data for problem enactment (QUALITY).
- I found managerial support for the extension of my research question to personal variables addressing CRM knowledge benefits and information management

competence. The feedback in relation to CRM knowledge benefits resulted in the latent construct customer relationship problem sense making enhancements (ENACENH) and the feedback in relation to information management competence resulted in the latent constructs customer relationship problem sense making information search modes (SCAN, FOCUS) and customer relationship problem sense making information search heuristics (REP, AVBLE, AA, POSI).

- There is a need to identify moderator variables that might impact QUALITY, ENACENH, SCAN, FOCUS, REP, AVBLE, AA, and POSI.

I presented these findings in [Abril 2001]. Although no major issues were found in the presentation, I received feedback about one of the comments from the participants in the research group. It was recommended to consider a manager's job function and CRM functions supported by the data warehouse as moderator variables in this research.

4.3.4 Questionnaire Design

This research phase covered the layout of the questionnaire and the content including the introduction. Guidelines in [Churchill Jr. 1979, 1991] were followed for developing the layout. The following sections were identified: Introduction, items about QUALITY, items about ENACENH, items about SCAN, FOCUS, REP, AVBLE, AA, and POSI, and finally items about the informant's personal data.

The production of the questionnaire followed an anchor and adjustment procedure. Up to four versions of the questionnaire were produced:

- Questionaire.v1: This was the first version of my questionnaire (a MS-word document). It came from my literature review (i.e. existing scales) and changes adapting them to the specific context, situation-within-context and unit of analysis of this research (see details in APPENDIX B: CONTENT VALIDITY RESEARCH FOCUS GROUP. QUESTIONNAIRE.v1). Questionnaire.v1 was input to the Content Validation Focus Group.
- Questionnaire.v2: It resulted from changes suggested in the Content Validation Focus Group. This version was implemented at Henley's web server (see details in

APPENDIX C: FACE VALIDITY RESEARCH FOCUS GROUP). Questionnaire.v2 had two different formats: an Acrobat PDF format and a HTML format. Both were input to the Face Validation Focus Group.

- Questionnaire.v3: It resulted from changes suggested in the Face Validation Focus Group. This version was implemented on Henley's web server in HTML format (see details in APPENDIX D: PILOT RESEARCH STUDY. QUESTIONNAIRE.v3). Questionnaire.v3 was input to the Pilot Study.
- Questionnaire.v4: It resulted from changes suggested in the Pilot Study. This version was implemented on Henley's web server in HTML format (see details in APPENDIX E: SURVEY QUESTIONNAIRE.v4). Questionnaire.v4 was input to the Survey Administration.

Selected items in the same construct were placed in a non-consecutive way, but always along with items using the same response scale format. Each item had a serial number. Bold fonts were used in order to highlight key aspects. Each item and explanation had at least one word in bold font.

The introduction section included Henley's logo, my name and e-mail address, my affiliation to Henley and the type of research to be conducted. Explanations about the areas to be covered by the items were short but enough directions were provided for the questionnaire. All the sections in the questionnaire (introduction included) were unchanged in the survey administration. This means that the entire sample responded to exactly the same questionnaire. However, salience [Schaefer and Dillman 1998] was addressed in the customized communications depending on the targeted group in the population (see Sample Framing and Data Collection Planning).

Explanations were inserted introducing groups of items, which sometimes included examples. The format of item-answer choices for the Likert-type scales was a table with the text for the items in the left column and the answer choices for the Likert-type scales to the right.

4.3.5 First and Fourth Operationalizations of First Order Latent Constructs

The eclectic nature of the theoretical basis in this research covering cognitive psychology (e.g. knowledge, self-efficacy, problem solving), behavior psychology (e.g. information search behavior), marketing (e.g. CRM), and managerial information processing (e.g. data quality) implies a cross-disciplinary set of measures.

The following is a comprehensive list of related operationalizations of this construct:

- Quality of marketing information as a 4 items 9-point Likert scale (1-Strongly disagree, 9- Strongly agree). 0.79 Cronbach alpha [Low and Mohr 2001]
- Researcher information quality-actionability 3 items 7-point semantic differential .87 Cronbach alpha [Moorman and Austin 1995]
- Researcher information quality-creativity 3 items 7-point semantic differential .87 Cronbach alpha [Moorman and Austin 1995]
- Researcher information quality-objectivity 5 items 7-point semantic differential .71 Cronbach alpha [Moorman and Austin 1995]
- Quality of information as a 6 items 7-point scale.(1-Strongly disagree, 7-Strongly agree) 0.76 Cronbach alpha [Goodhue and Thompson 1995]
- Quality of data warehouse data (general) 4 items 7-point Likert scale (1-strongly disagree, 7-strongly agree). 0.84 Cronbach alpha [Wixom and Watson 2001]. This scale measures the improvement in quality of the Data warehouse data versus the source systems which does not necessarily mean an acceptable level of quality.
- Quality of cross-functional marketing information (subscale of the cross-functional integration scale) in terms of accuracy, promptness, and timeliness of information exchanges as 3 items 7-point Likert scale (1-strongly disagree, 7-strongly agree). Reported studies with 0.69, 0.73, 0.79, and 0.80 Cronbach alphas [Song *et al.* 2000]. This scale does not measure information integration.
- Quality of data warehouse data (general) in terms of accuracy, format, and preciseness as a 7 items 5-point Likert scale (1-almost never, 2- some of the time, 3- about half of the time, 4- most of the time 5-almost always). No reported Cronbach alpha [Chen *et al.* 2000]- Quality-reliability (subscale of the

Information value scale) as a 5 items semantical differential scale, 0.83 Cronbach alpha [Swanson 1987]

- Quality of information as a 4 items 5-point Likert scale 0.81Cronbach alpha [Srinivasan 1985]
- Quality of information as a 10 items 7-point Likert scale 0.95 Cronbach alpha [Seddon and Kiew 1994]
- Quality of information as a 7 items 5-point Likert scale 0.89 Cronbach alpha [O'Reilly III 1982]
- Quality of information as a 4 items, two 7-point semantic differential scales per item 0.88 Cronbach alpha [Kettinger and Lee 1994]
- System's output quality as a 2 items, 7-point Likert (1-stronglu disagree, 2-moderatelyt disagree, 3-somewhat disagree, 4-neutral (neither disagree nor agree), 5-somewhat agree, 6-moderately agree, 7-strongly agree) ranged from 0.82 to 0.98) Cronbach alpha [Venkatesh and Davis 2000]
- Information value as a 7 items 5-point Likert scale (1-not at all, 3-neutral,5-totally) 0.81Cronbach alpha [Vandenbosch and Higgins 1995]
- Perceived source quality as a 2 items 5-point Likert scale no reported Cronbach alpha [Choo 1993]
- Perceived usefulness as a 6 items 5-point Likert scale (1-Low, 5-High) 0.96 Cronbach alpha [Vandenbosch and Huff 1997] and .98 [Davis 1989] 0.94 & 0.93 [Adams, Nelson and Todd 1992], and .94 & .96 [Hendrickson, Massey and Cronan 1993]
- Information relevance as a 3 items 5-point Likert scale (1-to a very great extent, 2-to a great extent, 3-to some extent, 4-to a little extent, 5-to a very little extent)). 0.89 Cronbach alpha [O'Reilly III 1982].
- Information relevance as a 2 items 7-point scale (1-Never, 4-A Fair Amount, 7- To a very great extent). 0.77 Cronbach alpha [Franz 1999]
- Problem solving utility as a 4 items 5-point Likert scale 0.81Cronbach alpha [Srinivasan 1985]

Unfortunately, a thorough analysis of these scales gives the following overall conclusions:

- It is very difficult to find a high order implementation with the exception of Moorman's researcher information quality scale [Moorman and Austin 1995].
- Most of the scales are too general. Some of them focus on the marketing function context which is too broad for the context of this research (i.e. CRM) and there are no situation-within-context (i.e. problem enactment) scales. The Srinivasan's problem solving utility scale [Srinivasan 1985] addresses a situation without a context.
- The term 'information' frequently refers to 'data'.

In relation to the high order latent construct problem sense making enhancements, the following is a comprehensive list of related operationalizations of this construct:

- Sense making as a 3 items 11-point Likert scale (0-Strongly disagree, 10-Strongly agree). 0.67 Cronbach alpha [Akgun 2001]
- Problem formulation process descriptors as a 14-items 8-point Likert scale (1-stronggly disagree, 8-strongly agree) No reported Cronbach alpha [Cowan 1988]. This scale focuses on four problem interpretations (i.e. strategic, operating, human, and technical)
- Enhancement of problem-solving ability as a 3 items 7-point Likert scale (1-Completly disagree, 7-Completely agree). 0.89 Cronbach alpha [Aldag and Power 1986]
- Problem statement adequacy as a 2 items 5-point Likert scale (1-Disagree completely, 5-Agree Completely). 0.74 Cronbach alpha [Aldag and Power 1986]
- Impact of change in current practice as a 3 items 7-point Likert scale (1-strongly disagree, 7-strongly agree) 0.81 Cronbach alpha [Lau and Woodman 1995]
- Intensity of change as a 4 items 7-point Likert scale (1-strongly disagree, 7-strongly agree) 0.80 Cronbach alpha [Lau and Woodman 1995]
- Individual Impact as a 3 items 7-point Likert scale (1-strongly disagree, 7-strongly agree) 0.88 Cronbach alpha [Spreitzer 1995]
- Team-level impact 6 items 0.93 Cronbach alpha [Kirkman and Rosen 1999] inspired in the scale of [Thomas and Tymon Jr.1993]

- Extent of changes caused by the DSS: as a 6 items 7-point scale (1-low, 7-high) 0.893 Cronbach alpha [Barki and Huff 1985]
- Self-efficacy as a 100 point scale, ranging in 10 unit intervals from 0 ("Cannot do"); through intermediate degrees of assurance, 50 ("Moderately certain can do"); to complete assurance, 100 ("Certain can do") [Bandura 2001, Pajares *et al.* 2001]
- General self-efficacy in vocational competence as a 17 item 14 point Likert scale (1-strongly disagree, 14-strongly agree). No reported Cronbach alpha [Sherer, Maddux, Mercadante, Prentice-Dunn, Jacobs and Rogers 1982]. This scale was re-examined by [Woodruff and Cashman 1993]
- Computer self-efficacy in executive support systems as a 5 items 5 point scale (1-not at all confident, 3-neutral, 5-Totally confident) 0.85 coefficient of internal consistency using Fornell's and Larker's [1981] measure [Vandenbosch and Higgins 1995]

Again, a thorough analysis of these scales gives the following overall conclusions:

- There are no high order implementations. Enactment mastery experience and enactment mastery expectancy are addressed by independent scales.
- Most of the scales are too general. They focus on a situation without a context.

In relation to the high order latent construct information search modes, it follows a comprehensive list of related operationalizations of this construct:

- Environmental scanning as a composite of frequency of scanning 7-items (one per sector) 6-point scale (0-never, yearly (1), quarterly (4), monthly (12), weekly (52), daily (365)), interest in scanning 7-items (one per sector) 5-point Likert scale (very low (1), very high (5)). No reported Cronbach alpha [Sawyerr *et al.* 2000]
- Scanning search 1 item 5-point scale (1-very seldom, 5-very often) [Gonzalez 2001]
- Focused search 1 item 5-point scale (1-very seldom, 5-very often) [Gonzalez 2001]
- Scanning search 4 items 7-point scale (1-never, 7-regularly) 0.929 Fornell and Larker alpha [Vandenbosch and Higgins 1996]

- Focused search 5 items 7-point scale (1-never, 7-regularly) 0.925 Fornell and Larker alpha [Vandenbosch and Higgins 1996]
- Exploratory research purpose as a 7 items 5-point Likert scale (1-Strongly agree, 2-Agree, 3-Neither agree nor disagree, 4-disagree,5- Strongly disagree) 0.76 Cronbach alpha [Deshpande and Zaltman 1982]
- Opportunity exploration as a 3 items 6-point scale (1-Never, 2-almost never, 3-sometimes, 4-fairly often, 5-very often 6-always) 0.71 Cronbach alpha [Kleysen and Street 2001]
- Scanning intensity 4 items (item #1 with a 5-point (1-Not difficult, 5-Extremely difficult), Item #4 with a 5-point (1-No extent,5-Great extent)) 0.68 Cronbach alpha [Maier *et al.* 1997]
- Issues scanned 5 item 5-point (1-No extent, 5-Great extent) 0.86 Cronbach alpha [Maier *et al.* 1997]
- Scanning method 1 item (1-Informally, 2-As needed by MIS function, 3-Regular schedule by MIS function, 4-As needed by team, 5-Regular schedule by team, 6-No response) [Maier *et al.* 1997]

The overall conclusion after analyzing these scales is that they are too general. They do not have a focus in situation-within-context.

In relation to the high order latent construct information search heuristics, the following is a comprehensive list of related operationalizations of this construct:

- Confirmatory research purpose as a 8 items 5-point Likert scale (1-Strongly agree, 2-Agree, 3-Neither agree nor disagree, 4-disagree, 5- Strongly disagree) 0.76 Cronbach alpha [Deshpande and Zaltman 1982]
- Incrementalism in 6 items Likert sub scale [Bailey and Johnson 1995]
- The ways of coping questionnaire (WCQ) [Folkman and Lazarus 1988] is a self-report instrument designed to assess situation-specific cognitions and actions. The respondent is instructed to focus on a particular episode during the past week that was experienced as either taxing or stressful and to respond to each of 150 items on a 4-point scale indicating the frequency with which a particular coping strategy was used (0 = does not apply or not used, 3 = used a great deal).

The respondent is instructed to focus on a particular episode during the past week that was experienced as either taxing or stressful and to respond to each of 150 items on a 4-point scale indicating the frequency with which a particular coping strategy was used (0 = does not apply or not used, 3 = used a great deal).

The overall conclusion after analyzing these scales it is that they are too general. They do not have a focus in situation-within-context.

Therefore, the operationalization of first order latent constructs had to follow the construct-oriented approach [Hough and Paullin 1994] as it follows:

- The theoretical basis underpinning the research model provided a set of well defined constructs. I took such constructs and reused the best available existing scales in the literature for such constructs.
- I introduced changes to the existing scales and adapted them to the specific context, situation-within-context and unit of analysis of this research (see CONTEXT, SITUATION AND UNIT OF ANALYSIS RESEARCH SPECIFICATIONS). I conducted focus groups in order to assess content and face validity.
- I used inductive analysis (e.g. [Hough and Paullin 1994, Hair Jr. *et al.* 1998]) in order to revise the scales.

As a consequence, this research required five successive operationalizations. Each of the first four is linked to a questionnaire version and the fifth and last one is linked to the subsequent analysis made after factor analysis (see details in APPENDIX F: DETAILS ON THE OPERATIONALIZATION OF VARIABLES). No new constructs were created, although, again, all the definitions were adapted to the specifics of context, situation-within-context and unit of analysis of this research. Details about operationalizations #1 and #4 of the first order latent environmental constructs in my research model can be seen in Table 4-5 for CUST, Table 4-6 for COMP, Table 4-7 for ENAC, Table 4-8 for INT, Table 4-9 for RICH, Table 4-10 for EQUI, and Table 4-11 for TRUST. Details about operationalizations #1 and #4 of the first order latent cognitive constructs in my research model can be seen in Table 4-12

CONSTRUCT	Customer insights
DEFINITION	Manager's perception on the extent to which the data warehouse customer relationship data is a source of customer insights (e.g. [Day 1994])
VARIABLE	CUST
(#) ITEMS in v1	(6) item_10.v1, item_11.v1, item_12.v1, item_13.v1, Item_14.v1, item_15.v1
RESPONSE FORMAT in v1	7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)
(#) ITEMS in v4	(6) item_1.v4, item_2.v4, item_3.v4, item_4.v4, Item_5.v4, item_6.v4
RESPONSE FORMAT in v4	Item_3.v4 and item_6.v4 reverse code. 7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)

Table 4-5 First and fourth operationalizations of CUST

For the operationalization of CUST (see Table 4-5) items item_10.v1, item_11.v1, item_12.v1, and item_13.v1 were adapted from the customer Orientation scale, 6 items 7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree), alpha 0.85 [Narver and Slater 1990]. Item_14.v1 was created in order to cover demographic data types [Shapiro and Bonoma 1984]. Item_15.v1 was created in order to cover behavioral data types [Peltier *et al.* 1998, Brewer and Richards 2001].

CONSTRUCT	Competitor insights
DEFINITION	Manager's perception on the extent to which the data warehouse customer relationship data is a source of competitor insights (e.g. [Day 1994])
VARIABLE	COMP
(#) ITEMS in v1	(2) item_16.v1, item_17.v1
RESPONSE FORMAT in v1	7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)
(#) ITEMS in v4	(3) item_7.v4, item_8.v4, item_9.v4
RESPONSE FORMAT in v4	7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)

Table 4-6 First and fourth operationalizations of COMP

For the operationalization of COMP (see Table 4-6) items item_16.v1, item_17.v1 and item_9.v4 were adapted from the Competitor Orientation scale, 4 items 7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree), alpha 0.71 [Narver and Slater 1990].

CONSTRUCT	Customer relationship problem enactment
DEFINITION	Manager's perception on the extent to which the data warehouse is a source of applicable and helpful customer relationship data for enacting customer relationship problems (e.g. [Fedorowicz and Lee 1998, Bailey and Pearson 1983, Venkatesh and Davis 2000])
VARIABLE	ENAC
(#) ITEMS in v1	(4) item_18.v1, item_19.v1, item_20.v1, item_21.v1
RESPONSE FORMAT in v1	Item_18.v1: 5-point (1- Not at all,2- A little extent, 3-Some extent, 4- Large extent, 5- Very relevant) Item_19.v1: 5-point scale (1-Never, 2-Rarely, 3-Sometimes,4-Frequently, 5-Most of the time) Item_20.v1 and item_21.v1: 7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)
(#) ITEMS in v4	(4) item_10.v4, item_11.v4, item_12.v4, item_13.v4
RESPONSE FORMAT in v4	Item_13.v4 reverse code. Item_10.v4: 5-point (1- Not at all,2- A little extent, 3-Some extent, 4- Large extent, 5- Very relevant) Item_11.v4: 5-point scale (1-Never, 2-Rarely, 3-Sometimes,4-Frequently, 5-Most of the time) Item_12.v4 and item_13.v4: 7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)

Table 4-7 First and fourth operationalizations of ENAC

For the operationalization of ENAC (see Table 4-7) items item_18.v1 and item_19.v1 were adapted from the Information Relevance scale, 2 items, one with 7-point Likert scale (1-Not at all, 4-A fair amount, 7-To a very great extent) and the other with 7-point (1-Never, 4-A Fair Amount, 7- To a very great extent), alpha 0.77 [Franz 1999].

Item_20.v1 was created in order to cover cognitive structure aspects (i.e. factors) in problem enactment [Sullivan and Weaver 2000]. Item_21.v1 v1 was created in order to cover cognitive structure (i.e. relationships) aspects in problem enactment [Sullivan and Weaver 2000].

CONSTRUCT	Data integration
DEFINITION	Manager's perception on the extent to which the data warehouse customer relationship data is normalized in terms of data definitions and logical data structures (e.g. [Goodhue *et al.* 1992, Goodhue *et al.* 2000, Peltier *et al.* 1998, Codd 1972])
VARIABLE	INT
(#) ITEMS in v1	(5) item_22.v1, item_23.v1, item_24.v1, item_25.v1, item_26.v1
RESPONSE FORMAT in v1	7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)
(#) ITEMS in v4	(3) item_14.v4, item_15.v4, item_16.v4
RESPONSE FORMAT in v4	7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)

Table 4-8 First and fourth operationalizations of INT

For the operationalization of INT (see Table 4-8) items item_25.v1 and item_26.v1 were adapted from the Data Warehouse Information integration scale, 3 items, 7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree), alpha 0.79 [Haley 1997]. Item_22.v1 was created in order to cover the redundancy/duplicates aspect of integration [Codd 1972]. Item_23.v1 v1 was created in order to cover the time-variance aspect of integration [Inmon 1996]. Item_24.v1 was created in order to cover the consistency aspect of integration [Codd 1972].

CONSTRUCT	Data richness
DEFINITION	Manager's perception on the extent to which the data warehouse customer relationship data is a source of customer relationship information (e.g. [Daft and Lengel 1984])
VARIABLE	RICH
(#) ITEMS in v1	(5) item_30.v1, item_32.v1, item_33.v1, item_34.v1, item_37.v1
RESPONSE FORMAT in v1	7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)
(#) ITEMS in v4	(5) item_20.v4, item_22.v4, item_23.v4, item_24.v4, item_27.v4
RESPONSE FORMAT in v4	Item_23.v4 reverse code 7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)

Table 4-9 First and fourth operationalizations of RICH

For the operationalization of RICH (see Table 4-9) all items were adapted from the Conceptual Utility for Product Strategy Development scale, 9 items, 7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree), alpha 0.80 [Moorman 1995].

CONSTRUCT	Data equivocality
DEFINITION	Manager's perception on the extent to which the data warehouse customer relationship data favors more than one interpretation for the enactment of customer relationship problems (e.g. [Goodhue *et al.* 2000, Jarke *et al.* 1999, Weick 1979, Daft and Weick 1984, Daft and Lengel 1986, Swanson 1987])
VARIABLE	EQUI
(#) ITEMS in v1	(3) item_27.v1, item_29.v1, item_36.v1
RESPONSE FORMAT in v1	Reverse code 7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)
(#) ITEMS in v4	(3) item_17.v4, item_19.v4, item_26.v4
RESPONSE FORMAT in v4	Reverse code 7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)

Table 4-10 First and fourth operationalizations of EQUI

For the operationalization of EQUI (see Table 4-10) all items were adapted from the equivocality scale, 3 items, 5-point scale (1-to a very great extent, 2-to a great extent, 3-to some extent, 4-to a little extent, 5-to a very little extent), alpha 0.74 [Brannick1998]. These are the same scales from Daft and Macintosh [1981] who reported a 0.73 Cronbach alpha. Another alternative scale initially reviewed but ultimately not considered was equivocality as a 2 items 7-point scale (1-Never, 4-A Fair Amount, 7- To a very great extent). 0.73 Cronbach alpha [Franz 1999].

CONSTRUCT	Data trustworthiness
DEFINITION	Manager's perception on the extent to which the data warehouse customer relationship data is regarded as true and credible evidence for the enactment of customer relationship problems (e.g. [Jarke *et al.* 1999])
VARIABLE	TRUST
(#) ITEMS in v1	(5) item_28.v1, item_31.v1, item_35.v1, item_38.v1, item_39.v1
RESPONSE FORMAT in v1	Item_28.v1 and item_31.v1 in reverse code. 7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)
(#) ITEMS in v4	(3) item_18.v4, item_21.v4, item_25.v4
RESPONSE FORMAT in v4	Item_21.v4 reverse code. 7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)

Table 4-11 First and fourth operationalizations of TRUST

For the operationalization of TRUST (see Table 4-11) item_28.v1, item_31.v1, and item_35.v1 were adapted from the Trust in the Marketing Information scale, 3 items, 7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree), alpha 0.81[Van Birgelen *et al.* 2000]. These authors adapted Moorman's 5 item scale from [Moorman *et al.* 1992]. Item_38.v1 was created in order to cover accuracy of data [Wang and Strong 1996, Huang *et al.* 1998, Rudra and Yeo 1999, Swanson 1987, Haley 1997]. Item_39.v1 was suggested in [O'Reilly III 1982]. Other alternative scales initially reviewed but ultimately not considered were:

- Information credibility as a 7 items scale No reported Cronbach alpha [Moenaert and Souder 1996].

- Trust in the marketing contact as a 6 items 5-point scale (1-strongly disagree 5-strongly agree) 0.77 Cronbach alpha [Maltz and Kohli 1996]
- Source credibility as a 13 items 5-point dichotomous scale 0.77 Cronbach alpha [Grooms 2001]

CONSTRUCT	Customer relationship problem declarative knowledge enhancement
DEFINITION	Manager's self-belief about the extent that knowledge on customer relationship problems has improved due to the data warehouse customer relationship data. Declarative knowledge is the set of stored situational cues and facts (e.g., types of customers and selling situations) which allows the manager to recognize and categorize customer relationship situations (e.g. [Porter and Inks 2000, Leidner *et al.* 1999, Campbell 1994])
VARIABLE	DKNOWE
(#) ITEMS in v1	(4) item_40.v1, item_41.v1, item_42.v1, item_43.v1
RESPONSE FORMAT in v1	7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)
(#) ITEMS in v4	(3) item_28.v4, item_29.v4, item_30.v4
RESPONSE FORMAT in v4	7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)

Table 4-12 First and fourth operationalizations of DKNOWE

For the operationalization of DKNOWE (see Table 4-12) all items were adapted from the Mental Model Enhancement scale, 4 items, 5-point (1- Not at all,2- A little extent, 3-Some extent, 4- Large extent, 5- Very relevant), alpha 0.90 [Leidner *et al.* 1999].

CONSTRUCT	Customer relationship problem enactment procedural knowledge enhancement
DEFINITION	Manager's self-belief about the extent that one's skills enacting customer relationship problems have improved due to the data warehouse customer relationship data. Procedural knowledge consists of routines, actions, strategies, or heuristics that apply to a task domain (e.g. [Porter and Inks 2000, Campbell 1994])
VARIABLE	PKNOWE
(#) ITEMS in v1	(4) item_44.v1, item_45.v1, item_46.v1, item_47.v1
RESPONSE FORMAT in v1	7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)
(#) ITEMS in v4	(4) item_31.v4, item_32.v4,item_33.v4,item_34.v4
RESPONSE FORMAT in v4	7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)

Table 4-13 First and fourth operationalizations of PKNOWE

For the operationalization of PKNOWE (see Table 4-13) all items were adapted from the procedural Rationality in the Strategic Decision-Making Process scale, 5 items, four of them in 7-point scale (1-Not at all, 4-A fair amount, 7-To a very great extent) and one in a 7-point reverse code (1-Mostly analytical, 7-Mostly intuitive), alpha 0.80 [Dean Jr. and Sharfman 1993].

CONSTRUCT	Customer relationship problem integrative complexity enhancement
DEFINITION	Manager's self-belief about the extent that the integrative complexity of the cognitive structures about customer relationship problems has improved due to data warehouse customer relationship data. Integrative complexity is the level of comprehensiveness - i.e. number of factors in the cognitive structure- and connectedness - i.e. links among the factors in the cognitive structure- (e.g. [McFadzean 1996, Wang and Chan 1995, Stone 1994])
VARIABLE	ICPLXE
(#) ITEMS in v1	(2) item_48.v1, item_49.v1
RESPONSE FORMAT in v1	7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)
(#) ITEMS in v4	(2) item_35.v4, item_36.v4
RESPONSE FORMAT in v4	7-point Likert scale (1-strongly disagree, 2-moderately disagree, 3-slightly disagree, 4-neither agree nor disagree, 5-slightly agree, 6-moderately agree, 7-strongly agree)

Table 4-14 First and fourth operationalizations of ICPLXE

For the operationalization of ICPLXE (see Table 4-14) all items were adapted from the Integrated Understanding scale, 4 items, 5-point (1- Not at all, 2- A little extent, 3-Some extent, 4- Large extent, 5- Very relevant), alpha 0.83 [Parker and Axtell 2001]. Other alternative scales initially reviewed but ultimately not considered were:

- Taxonomic complexity [Gonzalez 2001] adapted from [Hodgkinson and Johnson 1994]
- Goal-derived categories [Gonzalez 2001] adapted from Personal Goal Salience in [Ratneshwar, Pechmann and Shocker 1996]
- Task complexity as a multivariate of three types of task complexity: component, coordinative, and dynamic complexity [Wood 1986]

- Size of the model as the sum of all the model components (e.g. factors, relationships) [Bajaj 2000]
- Simple vs. complex explanations for others' behavior as a 4 items Likert scale. 0.61 Cronbach alpha [Porter and Inks 2000]
- Knowledge complexity as a 2 items x-point Likert scale 0.50 Cronbach alpha [Phillips 2001]
- Size = F + R [Bajaj 2000]
- Cognitive complexity as the ratio of connectedness to comprehensiveness within a problem interpretation [Sullivan and Weaver 2000]
- Integrated knowledge scale 0.62 Cronbach alpha [Phillips 2001]
- (Organizational) knowledge integration as a 4-items 7-point Likert scale (1 = strongly disagree, 7 = strongly agree) 0.8747 Cronbach alpha [Bontis and Fitz-enz 2002]

CONSTRUCT	Customer Relationship Problem Enacting Self-Efficacy
DEFINITION	The strength in the manager's self-belief in one's capabilities to execute given types of performances enacting prospective customer relationship problematic situations (e.g. [Bandura 1997, Bandura 1986])
VARIABLE	SELFE
(#) ITEMS in v1	(12) item_50.v1, item_51.v1, item_52.v1, item_53.v1, item_54.v1, item_55.v1, item_56.v1, item_57.v1, item_58.v1, item_59.v1, item_60.v1, and item_61.v1. Items item_50.v1 to item_53.v1 are the LSELFE subscale measuring self-efficacy with low challenging situations, items item_54.v1 to item_57.v1 are the MSELFE subscale measuring self-efficacy with moderately challenging situations, and items item_58.v1 to item_61.v1 are the HSELFE subscale measuring self-efficacy with very challenging situations.
RESPONSE FORMAT in v1	11-point scale (0-No chance at all, 5: A 50/50 10- Completely certain)
(#) ITEMS in v4	(12) item_37.v4, item_38.v4, item_39.v4, item_40.v4, item_41.v4, item_42.v4, item_43.v4, item_44.v4, item_45.v4, item_46.v4, item_47.v4, and item_48.v4. Items item_37.v4 to item_40.v4 are the LSELFE subscale measuring self-efficacy with low challenging situations, items item_41.v4 to item_44.v4 are the MSELFE subscale measuring self-efficacy with moderately challenging situations, and items item_45.v4 to item_48.v4 are the HSELFE subscale measuring self-efficacy with very challenging situations.
RESPONSE FORMAT in v4	11-point scale (0-No chance at all, 5: A 50/50 chance, 10-Completely certain)

Table 4-15 First and fourth operationalizations of SELFE

For the operationalization of SELFE (see Table 4-15) all items were adapted from the Group Problem Solving Self-Efficacy scale, 12 items, 100 point scale ranging in 5 unit intervals (0-no chance at all, 25-A slight chance, 50-A 50-50 Chance, 75- A good chance, 100-completely certain), alpha 0.95 [Wolf 1997].

CONSTRUCT	Intensity of Data Warehouse Scanning Search
DEFINITION	Manager's perceptions on the personal's amount of effort scanning the data warehouse making sense of customer relationship problems. Scanning search is the proactive and exploratory information search behavior mode people exhibit when they browse through information without a particular problem to solve (e.g. [Maier *et al.* 1997, Boynton 1987])
VARIABLE	SCAN
(#) ITEMS in v1	(4) item_62.v1, item_63v1, item_64.v1, item_67.v1
RESPONSE FORMAT in v1	Items item_62.v1, item_63v1, and item_64.v1:5-point scale (1-Never, 2-Rarely, 3-Sometimes,4-Frequently, 5-Most of the time) Item_67.v1: 100- sum
(#) ITEMS in v4	(2) item_49i.v4,item_50i.v4
RESPONSE FORMAT in v4	100-point constant sum

Table 4-16 First and fourth operationalizations of SCAN

For the operationalization of SCAN (see Table 4-16) items item_62.v1, item_63.v1 and item_67.v1 were adapted from the Managerial Scanning scale, 3 items, one item in 5 point scale (1-Very rarely, 2-Infrequently, 3-Occasionally, 4-Frequently, 5-Almost always), one item in 5 point scale (1-Less than 2 hr/week, 2-About 3 Hours/week, 3-About 4 Hours/week, 4-About 1-2 Hours/Day, 5-More than 3 hours/Day), one item in 5-point (1- Not at all, 2- A little extent, 3-Some extent, 4-Large extent, 5- Very relevant), alpha 0.73 [Boynton 1987].Item_64.v1 was created in order to cover new problems/issues as a predictor [Deshpande and Zaltman 1982].

CONSTRUCT	Intensity of Data Warehouse Focused Search
DEFINITION	Manager's perceptions on the personal's amount of effort in focused search on the data warehouse making sense of customer relationship problems. Focused search is the reactive and directed information search behavior mode people exhibit when they are looking for information specific to a problem to be addressed or question to be answered (e.g. [Maier *et al.* 1997, Boynton 1987])
VARIABLE	FOCUS
(#) ITEMS in v1	(3) item_65.v1, item_66.v1, item_68.v1
RESPONSE FORMAT in v1	Items item_65.v1, and item_66.v1: 5-point scale (1-Never, 2-Rarely, 3-Sometimes, 4-Frequently, 5-Most of the time) Item_68.v1: 100-point constant sum
(#) ITEMS in v4	(2) item_49ii.v4, item_50ii.v4
RESPONSE FORMAT in v4	100-point constant sum

Table 4-17 First and fourth operationalizations of FOCUS

For the operationalization of FOCUS (see Table 4-17) all items were adapted from the Managerial Scanning scale, 3 items, one item in 5 point scale (1-Very rarely, 2-Infrequently, 3-Occasionally, 4-Frequently, 5-Almost always), one item in 5 point scale (1-Less than 2 hr/week, 2-About 3 Hours/week, 3-About 4 Hours/week, 4-About 1-2 Hours/Day, 5-More than 3 hours/Day), one item in 5-point (1- Not at all,2-A little extent, 3-Some extent, 4- Large extent, 5- Very relevant), alpha 0.73 [Boynton 1987].

CONSTRUCT	Intensity of the Data Warehouse Representativeness Search Heuristic
DEFINITION	Manager's perceptions on the personal's amount of effort using representativeness heuristic searching the data warehouse in order to make sense of customer relationship problems. Representativenes heuristic refers (i) To assess the probability of a situation as representative of a category (e.g. a 'price-lowering by a competitor' situation can have common information with an 'attempt action to gain market-share' pattern), or (ii) to make generalizations based on new information about a sample - i.e. the sample is representative of a large population- (e.g. to assess the national market success of a new product line based on the data likehood ratio of a test market). This heuristic implies that the search ends when there is a satisficing fit between information about a situation and information about a category (e.g. [Chi and Fan 1997, Wright 1980])
VARIABLE	REP
(#) ITEMS in v1	(3) item_69.v1, item_70.v1, item_77.v1
RESPONSE FORMAT in v1	Items item_69.v1, and item_70.v1:5-point scale (1-Never, 2-Rarely, 3-Sometimes, 4-Frequently, 5-Most of the time) Item_77.v1: 100-point constant sum
(#) ITEMS in v4	(4) item_51i.v4, item_52i.v4, item_53i.v4, item_54i.v4
RESPONSE FORMAT in v4	100-point constant sum

Table 4-18 First and fourth operationalizations of REP

For the operationalization of REP (see Table 4-18) all items were adapted from the Managerial Scanning scale, 3 items, one item in 5 point scale (1-Very rarely, 2-Infrequently, 3-Occasionally, 4-Frequently, 5-Almost always), one item in 5 point scale (1-Less than 2 hr/week, 2-About 3 Hours/week, 3-About 4 Hours/week, 4-About 1-2 Hours/Day, 5-More than 3 hours/Day), one item in 5-point (1- Not at all, 2- A little extent, 3-Some extent, 4- Large extent, 5- Very relevant), alpha 0.73 [Boynton 1987].

CONSTRUCT	Intensity of the Data Warehouse Availability Search Heuristic
DEFINITION	Manager's perceptions on the personal's amount of effort using availability heuristic searching the data warehouse in order to make sense of customer relationship problems. Availability heuristic refers to assess the probability of a situation as a function of prior situations (e.g. a marketer considering a series of occurrences of actual costs incurred in past editions of a campaign when estimating the cost for a new edition of such campaign). This heuristic implies the search for easily accessible information about relevant precedents. The search ends once recent, salient information about a relevant precedent is found (e.g. [Chi and Fan 1997, Wright 1980])
VARIABLE	AVBLE
(#) ITEMS in v1	(3) item_71.v1, item_72.v1, item_78.v1
RESPONSE FORMAT in v1	Items item_71.v1, and item_72.v1: 5-point scale (1-Never, 2-Rarely, 3-Sometimes, 4-Frequently, 5-Most of the time) Item_78.v1: 100-point constant sum
(#) ITEMS in v4	(4) item_51ii.v4, item_52ii.v4, item_53ii.v4, item_54ii.v4
RESPONSE FORMAT in v4	100-point constant sum

Table 4-19 First and fourth operationalizations of AVBLE

For the operationalization of AVBLE (see Table 4-19) all items were adapted from the Managerial Scanning scale, 3 items, one item in 5 point scale (1-Very rarely, 2-Infrequently, 3-Occasionally, 4-Frequently, 5-Almost always), one item in 5 point scale (1-Less than 2 hr/week, 2-About 3 Hours/week, 3-About 4 Hours/week, 4-About 1-2 Hours/Day, 5-More than 3 hours/Day), one item in 5-point (1- Not at all, 2- A little extent, 3-Some extent, 4- Large extent, 5- Very relevant), alpha 0.73 [Boynton 1987].

CONSTRUCT	Intensity of the Data Warehouse Anchoring and Adjustment Search Heuristic
DEFINITION	Manager's perceptions on the personal's amount of effort using anchoring and adjustment heuristic searching the data warehouse in order to make sense of customer relationship problems. Anchoring and adjustment heuristic refers essentially to the trial and error method (e.g. a marketer setting the price of a product starting with a baseline price and making a number of impact analysis in several of the cost components). This heuristic implies a recursive process and each step involves a search for additional information and an adjustment of the previous assessment. The search ends once the adjustments are not improving an implicit/explicit value function on the information found (e.g. [Chi and Fan 1997])
VARIABLE	AA
(#) ITEMS in v1	(3) item_73.v1, item_74.v1, item_79.v1
RESPONSE FORMAT in v1	Items item_73.v1, and item_74.v1:5-point scale (1-Never, 2-Rarely, 3-Sometimes, 4-Frequently, 5-Most of the time) Item_79.v1: 100-point constant sum
(#) ITEMS in v4	(4) item_51iii.v4, item_52iii.v4, item_53iii.v4, item_54iii.v4
RESPONSE FORMAT in v4	100-point constant sum

Table 4-20 First and fourth operationalizations of AA

For the operationalization of AA (see Table 4-20) all items were adapted from the Managerial Scanning scale, 3 items, one item in 5 point scale (1-Very rarely, 2-Infrequently, 3-Occasionally, 4-Frequently, 5-Almost always), one item in 5 point scale (1-Less than 2 hr/week, 2-About 3 Hours/week, 3-About 4 Hours/week, 4-About 1-2 Hours/Day, 5-More than 3 hours/Day), one item in 5-point (1- Not at all, 2- A little extent, 3-Some extent, 4- Large extent, 5- Very relevant), alpha 0.73 [Boynton 1987].

CONSTRUCT	Intensity of the Data Warehouse Positivy Search Heuristic
DEFINITION	Manager's perceptions on the personal's amount of effort using positivy heuristic searching the data warehouse in order to make sense of customer relationship problems. Positivy heuristic refers to confirm the probability of a situation using the trial and error method (e.g. a marketer looking for issues in accounts that were predicted to have a high risk). This heuristic implies a search for information that is fundamentally consistent with existing beliefs, theories and cognition. The search ends once the information found confirms the probability of a situation (e.g. [Evans 1989])
VARIABLE	POSI
(#) ITEMS in v1	(3) item_75.v1, item_76.v1, item_80.v1
RESPONSE FORMAT in v1	Items item_73.v1, and item_74.v1:5-point scale (1-Never, 2-Rarely, 3-Sometimes, 4-Frequently, 5-Most of the time) Item_79.v1: 100-point constant sum
(#) ITEMS in v4	(4) item_51iv.v4, item_52iv.v4, item_53iv.v4, item_54iv.v4
RESPONSE FORMAT in v4	100-point constant sum

Table 4-21 First and fourth operationalizations of POSI

For the operationalization of POSI (see Table 4-21) all items were adapted from the Managerial Scanning scale, 3 items, one item in 5 point scale (1-Very rarely, 2-Infrequently, 3-Occasionally, 4-Frequently, 5-Almost always), one item in 5 point scale (1-Less than 2 hr/week, 2-About 3 Hours/week, 3-About 4 Hours/week, 4-About 1-2 Hours/Day, 5-More than 3 hours/Day), one item in 5-point (1- Not at all, 2- A little extent, 3-Some extent, 4- Large extent, 5- Very relevant), alpha 0.73 [Boynton 1987].

4.3.6 Operationalization of Moderator Variables and Qualification Requirements

Details about operationalizations #1 and #4 of the moderator variables can be seen in Table 4-22 for JOBFUNC and in Table 4-23 for DWFUNC.

CONSTRUCT	CRM job function
DEFINITION	CRM job function primarily performed by the informant. Four CRM job functions were studied: Sales support, marketing support, customer services support and other CRM function
VARIABLE	JOBFUNC
(#) ITEMS in v1	(1) item_7.v1
RESPONSE FORMAT in v1	Single choice 4-categories (1.- Sales support , 2.- Marketing support, 3.- Customer service support, 4.- Other CRM functions)
(#) ITEMS in v4	(1) item_61v.v4
RESPONSE FORMAT in v4	Single choice 4-categories (1.- Sales support , 2.- Marketing support, 3.- Customer service support, 4.- Other CRM functions)

Table 4-22 First and fourth operationalizations of JOBFUNC

CONSTRUCT	CRM data warehouse Function
DEFINITION	CRM functions supported by the data warehouse. Four CRM job function s were studied: Sales support, marketing support, customer services support and Other CRM function
VARIABLE	DWFUNC
(#) ITEMS in v1	(1) item_6.v1
RESPONSE FORMAT in v1	Multiple choice 4-categories (1.- Sales support , 2.- Marketing support, 3.- customer service support, 4.- Other CRM support)
(#) ITEMS in v4	(1) item_60v.v4
RESPONSE FORMAT in v4	Multiple choice 4-categories (1.- Sales support , 2.- Marketing support, 3.- Customer service support, 4.- Other CRM functions)

Table 4-23 First and fourth operationalizations of DWFUNC

Details about operationalizations #1 and #4 of the qualification variables can be seen in Table 4-24 for time enacting CRM problems, Table 4-25 for time searching/requesting customer relationship information, Table 4-26 for time that the informant has been engaged in CRM tasks, Table 4-27 for time that the informant has been an user of the firm's data warehouse customer data, and Table 4-28 for time the firms' data warehouse is supporting CRM

REQUIREMENT	Time enacting CRM problems
DEFINITION	Manager's perceptions on the time spent enacting CRM problems
(#) ITEMS in v1	(1) item_4.v1
RESPONSE FORMAT in v1	11-point scale (1-0% No time at all, 2-10%, 3-20%, 4-30%, 5-40%, 6-50%, 7-60%, 8-70%, 9-80%, 10-90%, 11-100% all my time)
(#) ITEMS in v4	(1) item_58.v4
RESPONSE FORMAT in v4	11-point scale (1-0% No time at all, 2-10%, 3-20%, 4-30%, 5-40%, 6-50%, 7-60%, 8-70%, 9-80%, 10-90%, 11-100% all my time)

Table 4-24 First and fourth operationalizations of time enacting CRM problems

REQUIREMENT	Time searching/requesting customer relationship information
DEFINITION	Manager's perceptions on the time searching/requesting customer relationship information in the firm's data warehouse
(#) ITEMS in v1	(1) item_5.v1
RESPONSE FORMAT in v1	11-point scale (1-0% No time at all, 2-10%, 3-20%, 4-30%, 5-40%, 6-50%, 7-60%, 8-70%, 9-80%, 10-90%, 11-100% all my time)
(#) ITEMS in v4	(1) item_59.v4
RESPONSE FORMAT in v4	11-point scale (1-0% No time at all, 2-10%, 3-20%, 4-30%, 5-40%, 6-50%, 7-60%, 8-70%, 9-80%, 10-90%, 11-100% all my time)

Table 4-25 First and fourth operationalizations of time searching/requesting customer relationship information

REQUIREMENT	Informant's CRM experience
DEFINITION	Time that the informant has been engaged in CRM tasks
(#) ITEMS in v1	(1) item_1.v1
RESPONSE FORMAT in v1	5-point scale (1-Less than 6 months 2-Between 6 months and 1 year, 3- Between 1 and 3 years, 4-Between 3 and 5 years, 5-More than 5 years)
(#) ITEMS in v4	(1) item_55.v4
RESPONSE FORMAT in v4	5-point scale (1-Less than 6 months 2-Between 6 months and 1 year, 3- Between 1 and 3 years, 4-Between 3 and 5 years, 5-More than 5 years)

Table 4-26 First and fourth operationalizations of time that the informant has been engaged in CRM tasks

REQUIREMENT	Informant's experience using firm's data warehouse customer data
DEFINITION	Time that the informant has been an user of the firm's data warehouse customer data
(#) ITEMS in v1	(1) item_2.v1
RESPONSE FORMAT in v1	5-point scale (1-Less than 6 months 2-Between 6 months and 1 year, 3- Between 1 and 3 years, 4-Between 3 and 5 years, 5-More than 5 years)
(#) ITEMS in v4	(1) item_56.v4
RESPONSE FORMAT in v4	5-point scale (1-Less than 6 months 2-Between 6 months and 1 year, 3- Between 1 and 3 years, 4-Between 3 and 5 years, 5-More than 5 years)

Table 4-27 First and fourth operationalizations of time that the informant has been an user of the firm's data warehouse customer data

REQUIREMENT	Time the firms' data warehouse is supporting CRM
DEFINITION	Time that firms' data warehouse has been operational supporting CRM
(#) ITEMS in v1	(1) item_3.v1
RESPONSE FORMAT in v1	5-point scale (1-Less than 6 months 2-Between 6 months and 1 year, 3- Between 1 and 3 years, 4-Between 3 and 5 years, 5-More than 5 years)
(#) ITEMS in v4	(1) item_57.v4
RESPONSE FORMAT in v4	5-point scale (1-Less than 6 months 2-Between 6 months and 1 year, 3- Between 1 and 3 years, 4-Between 3 and 5 years, 5-More than 5 years)

Table 4-28 First and fourth operationalizations of time the firms' data warehouse is supporting CRM

4.3.7 Content Validation Focus Group

Once I concluded the first version of my questionnaire, I conducted a moderated e-mail focus group research (e.g. [Adriaenssens and Cadman 1999, Curasi 2001, DeLorme, Zinkhan and French 2001]). Details of this exploratory research are included in APPENDIX B: CONTENT VALIDITY RESEARCH FOCUS GROUP. QUESTIONNAIRE.v1. Again, I followed the recommended guidelines of Stewart and Shamdasani [1998] of planning, selecting participants, moderating, data collection and principles of analysis.

This focus group provided great value in terms of weaknesses and risks to content validity of the scales in my questionnaire. Weakness W1 (i.e. self-efficacy low difficulty level subscale was not consistent with the other two subscales) was addressed in correcting items #50 to #53. Weakness W2 (i.e. most of the items were positively loaded) was addressed (i) introducing a 30% of reversed items and (ii) at least a reversed item per construct. It needs to be noted that this subject (i.e. reversed items) is controversial as there are researchers who argue that reverse items introduce error in data collection because of informants' mistakes understanding the wording. In fact, this is my experience: Most of the reversed items were dropped after the pilot.

Weakness W3 (i.e. vague response format for scales in Part III) was addressed by changing the response format in all the scales in Part III to a constant sum.

Risks R1 (i.e. few items in some scales) and R3 (i.e. long questionnaire) were kept until testing of unidimensionality of items. Risk R2 (i.e. lack of qualitative views) was addressed by the planned qualitative interventions along the research project. Implemented changes in the scales originated in a second version of the questionnaire.

4.3.8 Web Questionnaire Design and Implementation

This research phase covered the infrastructure supporting the sample data collection. The second version of the questionnaire was implemented as a web questionnaire [Solomon 2001] by Karen Whiting, an administrator at Henley, using the TeleForm software. Two formats of the second version of the questionnaire were implemented: An Acrobat PDF web questionnaire and a HTML web questionnaire. Something to be considered in this research phase was that most of the communications in the survey administration phase should be done by e-mails and that attachments could be perceived as infected e-mails. Therefore, one of my requirements was that e-mails should be sent with no attachments what so ever.

Sample surveys are subject to four major sources of error which need to be addressed to in order to have confidence in the precision of the sample survey estimates. These errors are [Dillman and Bowker 2001]:

- Coverage error: The result of all units in a defined population not having a known nonzero probability of being included in the sample drawn to represent the population.
- Sampling error: The result of surveying a sample of the population rather than the entire population.
- Measurement error: The result of inaccurate responses that stem from poor question wording, poor interviewing, survey mode effects and/or some aspect of the respondent's behavior.

- Nonresponse error: The result of nonresponse from people in the sample, who, if they had responded, would have provided different answers to the survey questions than those who did respond to the survey.

Dillman and Bowker [2001] suggested fourteen design principles for web questionnaires that mitigate such errors. Table 4-29 illustrates the compliance of the web questionnaire with these principles and the mitigated source of error. It shows that except for principle #2 (PIN) all of them were implemented.

4.3.9 Face Validation Focus Group

As reported in the Web Questionnaire Design section, the second version of my questionnaire was implemented in two formats: An Acrobat PDF format and a HTML format; both were accessible via two URLs to Henley's web server. I conducted a moderated e-mail focus group research (e.g. [Adriaenssens and Cadman 1999, Curasi 2001, DeLorme *et al.* 2001]) with the main objective of exploring the scales' face validity. Additionally, this research represented a great opportunity for receiving feedback in relation to the questionnaire design. Details of this exploratory research are included in APPENDIX C: FACE VALIDITY RESEARCH FOCUS GROUP. I followed the recommended guidelines of Stewart and Shamdasani [1998] of planning, selecting participants, moderating, data collection and principles of analysis.

The focus group clearly recommended that the HTML format was superior and that it would be confusing to provide a choice HTML vs. PDF to targeted informants. Therefore, the PDF format was abandoned keeping the HTML format as the single instrument for my sample data collection. The following is a summary of the changes implemented:

- Except items in the respondent's personal data section, all the other questions were changed to 'optional'. As reported in the Web Questionnaire Design section, the second version of the questionnaire rejected submissions with empty responses. This could increase error by forcing the informant to make undesirable responses and/or random responses just to submit the questionnaire as recognized in weaknesses W4

(i.e. once a response was provided you could not leave it blank) and W5 (i.e. all responses were mandatory). Obviously, my corrective action increased the 'blank response' risk.

- Diversity of scales (e.g. 5-point Likert, 7-point Likert) was reduced. The focus group considered diversity of scales (weaknesses W6, W8, and W9) to be confusing and tiring from an informant perspective. With this feedback, I considered that I had enough justification for changing the response format of some scales.
- Overall, the focus group assessed the questionnaire as 'demanding' and too long (risk W1). This finding was consistent with the feedback reported in APPENDIX B: CONTENT VALIDITY RESEARCH FOCUS GROUP. QUESTIONNAIRE.v1. Corrective action was taken by (i) reducing demographic items required for validating qualified informants, and (ii) working over the layout (e.g. fonts, page setup).
- Wording was the subject of criticisms in several items. The focus group considered that the code reverse was very confusing and recommended changing them to more straightforward language. This input was conflicting with the corrective action implemented as part of the findings reported in APPENDIX B: CONTENT VALIDITY RESEARCH FOCUS GROUP. QUESTIONNAIRE.v1. I decided to keep the existing code reverse items. A few items were reworded in order to make the language more understandable. For example, terms like 'enactment', 'a-priori', and 'heuristics' changed to 'understanding', 'preliminary', and 'strategy'.

Design Principle	Error Types				Implementation in Web Questionnaire
	Sampling	Coverage	Measurement	Non-response	
1. Introduce the web questionnaire with a welcome screen that is motivational, emphasizes the ease of responding, and instructs respondents on the action needed for proceeding to the next page.				X	Implemented.
2. Provide a PIN number for limiting access only to people in the sample.	X	X			Feature not supported by
3. Choose for the first question an item that is likely to be interesting to most respondents, easily answered, and fully visible on the first screen of the questionnaire.				X	Implemented.
4. Present each question in a conventional format similar to that normally used on paper self-administered questionnaires.			X	X	Implemented.
5. Restrain the use of color so that figure/ground consistency and read-ability are maintained, navigational flow is unimpeded, and measurement properties of questions are maintained.			X		Implemented.
6. Avoid differences in the visual appearance of questions that result from different screen configurations, operating systems, browsers, partial screen displays and wrap-around text.		X	X	X	Implemented.
7. Provide specific instructions on how to take each necessary computer action for responding to the questionnaire and other necessary instructions at the point where they are needed.				X	Implemented.
8. Use drop-down boxes sparingly, consider the mode implications, and identify each with a "click here" instruction.			X		Drop-down boxes were not used.
9. Do not require respondents to provide an answer to each question before being allowed to answer any subsequent ones.				X	Implemented.
10. Provide skip directions in a way that encourages marking of answers and being able to click to the next applicable question.			X		There were no "go to" directions.
11. Construct web questionnaires so they scroll from question to question unless order effects are a major concern, and/or telephone and web questionnaire results are being combined.		X	X	X	Implemented.
12. When the number of answer choices exceeds the number that can be displayed in a single column on one screen, consider double-banking with an appropriate grouping device to link them together.			X		Answer choices per item fit in the screen.
13. Use graphical symbols or words that convey a sense of where the respondent is in the completion process, but avoid ones that require significant increases in computer memory.		X		X	Implemented.
14. Exercise restraint in the use of question structures that have known measurement problems on paper questionnaires, e.g., check-all that-apply and open-ended questions.			X	X	Implemented.

Table 4-29 Compliance with web questionnaire design principles

- Sequence of the four items #51 to # 54 in the second version of the questionnaire was considered confusing (risk W8). The sequence of the topics of these items was 'relative frequencies for items A' & 'relative frequencies for items B' & 'time for items A' & 'time for items B'. I changed this sequence to 'relative frequencies for items A' & 'time for items A' & 'relative frequencies for items B' & 'time for items B'.
- Statements protecting my rights as intellectual owner were criticized (risks W12 and W13). I did not remove them.
- The language in the introduction section in the questionnaire included a message "To help marketers..." assuming that this would encourage informants to complete the questionnaire. The focus group (risk W9) recommended highlighting the doctoral aspect of this research. I replaced the criticized language by the language "you are contributing to this doctoral research".

Implemented changes created a third version of the questionnaire.

4.3.10 Sample Framing and Data Collection Planning

One important aspect in judging a sample as typical of the population as a whole is that the sample has the same variability as that of the population (e.g. [Curwin and Slater 1996]). In simple terms, the sample should include all particular types of informants and all the informants should have the same chance of being selected. The population of this research is defined by the unit of analysis "the individual manager engaged in customer relationship processes (i.e. the context) understanding customer relationship problems (i.e. the situation-within-context)" who satisfies a clear profile in terms of experience, actual dedication, and context (see ORGANISATION OF THE STUDY). Numerous research studies (e.g. [Muller 2002, Gottschalk 1998]) have suggested professional profiles for the population consisting of the affiliation to professional associations, affinity to specific professional subjects, subscription to specialized publications, and attendance of specialized conferences. It seems reasonable to assume that the population in this research meets the profile described in Table 4-30. Therefore, the following step was to identify sources of prospective informants with this profile. This approach leads to a judgmental sample [Churchill Jr. 1991].

[Individual managers (i) subscribers of CRM/data warehouse related publications AND (ii) working for organizations that are customers of data warehouse database management systems providers]

OR

[Individual managers (i) members of CRM/data warehouse related associations AND (ii) working for organizations that are customers of data warehouse database management systems providers]

OR

[Individual managers (i) participating in CRM/data warehouse related forums AND (ii) working for organizations that are customers of data warehouse database management systems providers]

OR

[Individual managers (i) attending CRM/data warehouse related conferences AND (ii) working for organizations that are customers of data warehouse database management systems providers]

Table 4-30 Profile of the informant

Experts in marketing regularly rely on non-probability sampling when conducting surveys (see [FJC 2000] p. 231.). Indeed, non-probability sampling is used widely in marketing research and the results of these studies are used by major companies in making decisions of considerable consequence (see [FJC 2000], p. 244.). I argue that I have not introduced any bias/criteria that might impact the representativeness of the variability of the population in this sample. For example, I have avoided the following biases: addressing data warehouses of just one provider, addressing data warehouses of a certain level (e.g. +1 terabyte) of customer raw data volume, addressing informants in the same country, and addressing informants in just one industry. The informant's profile in Table 4-30 suggests the following types of sources of prospective informants to be addressed in order to obtain my sample:

- CRM and data warehouse related publications
- CRM and data warehouse related associations
- CRM and data warehouse forums
- CRM and data warehouse conferences

I generated contact details (leads) from the identified sources. In general, a lead consisted of the following details: name, position, organization, e-mail address, telephone, and regular mail address. Position and regular mail address were not always available. All the leads were classified with respect to their positions as stated in Table 4-31. Disclosed positions were relevant because respondents seem more likely to respond to things concerning a salient experience than they are to things concerning a nonsalient or less interesting and/or memorable experience [Dillman and Carley-Baxter 2000]. An experience (e.g. answering a questionnaire) is salient for an individual when his/her attention is held by the experience and he/she is able to make inferences about the experience (e.g. [Dillman and Carley-Baxter 2000]).

Code	Position
11	DWH/CRM Customer (DB Administrator)
12	DWH/CRM Customer (DWH/IT Executive -IT architecture-)
13	DWH/CRM Customer (Data design-DW architecture- & sourcing)
21	DWH/CRM Customer (End-user)
22	DWH/CRM Customer (Functional executive & BI -DSS, DM-)
C	DWH/CRM Customer (no disclosed position)
31	Data Base/DWH User Group Representative

Table 4-31 Types of contacts and types of sources of prospective informants

A-priori (i.e. before checking the criteria for qualified informants) codes #21 and #22 were the positions considered as leads to prospective informants (my unit of analysis). Therefore, two types of leads emerged (see Table 4-32) and they should be approached in different ways: Direct leads and agentic leads. A direct lead was formed by contact details to prospective informants. In the case of conferences, this data came from the disclosed registration data by participants. Direct leads from forums required a tedious manual sequential scan of all (sic.) the e-mail communications taking place in such forums, looking for disclosed contact data. Under the category of agentic leads are disclosed positions #11, #12, #13, C, and #31. The idea of agentic leads was to approach individuals and request them to act as my research agent by forwarding my invitation to qualified individuals (the criteria for qualifications were clearly stated). End users groups were associations from which I extracted agentic leads (code #31) by accessing to the disclosed contact details of

their contact members in their web-portals. The support of agentic leads had two shapes: (i) Passive promotion in their web-portals with a text describing this research and a URL to the questionnaire and (ii) active promotion contacting their members by e-mail inviting them to participate.

		CONTACT TYPE	
		Agentic Lead	Direct Lead
SOURCE TYPE	Conferences	11&12&C, 13	22&21
	Forums	11&12&C, 13	22&21
	Associations	31	
	Publications	31	

Table 4-32 Types of contacts and types of sources of prospective informants

The invitation to participate in the research, in addition to salient information, should include a stimulus because past research shows that such a combination has significant influence on response rate. [Dillman and Carley-Baxter 2000]. I created four segments (see Table 4-33) and each one was addressed with a unique combination salient information and stimulus:

- Segment 22&21: Direct leads to prospective informants formed by end users, functional executives (e.g. marketing, sales, customer support), business intelligence practitioners, decision support systems practitioners, and data mining practitioners.
- Segment 11&12&C: Agentic leads formed by information technology individuals and undisclosed positions.
- Segment 13: Agentic leads formed by data integration practitioners.
- Segment 31: Agentic leads formed by end users groups' representatives.

This implied to (i) segment the available leads, (ii) design the appropriate language to be used in the communications (e-mails and letters) per segment, and (iii) identify the sequence when more than one lead was available in the same organization. Calls for participation/support were designed following recommendations from [Fahy 1998] and [Jobber and O'Reilly 1998] in order to increase response rate. Only segment 21&22 received calls for participation. The other three segments were addressed with

calls for agentic support (see a few samples in APPENDIX G: CALLS FOR PARTICIPATION SUPPORT).

SEGMENT	SALIENT INFORMATION	STIMULUS
21&22	Data Warehouse as a source of data for informant's job Struggle understanding CRM problems	Access to summary of results Licensing Questionnaire usage Benchmark
11&12&C	End users' satisfaction Data Warehouse contributes with value to the firm	Access to summary of results Licensing Questionnaire usage Benchmark
13	Data quality in the Data Warehouse Data Warehouse contributes with value to the firm	Access to summary of results Licensing Questionnaire usage Benchmark
31	Data quality in the Data Warehouse	Involvement promoting research

Table 4-33 Salient information and stimulus per segment

I collected contact details of 21 associations and publications and over 5,500 leads to individuals which required intensive data cleaning (e.g. identify duplicates, replace blanks by data available from other leads – for example mail address for the organization, inferring the e-mail address when I knew the domain). After data cleaning, I had 3,517 leads to individuals with the following procedences (i) 1,360 leads were extracted from forum in DataWarehouse.com, forum in Datawarehousing.com, forum in The Data Warehouse Institute, forum in CRMguru.com, forum in IntelligentCRM.com, forum in SearchCRM.com, DB2 users group, Oracle users group, Teradata users group, and SQL Server users group, and (ii) 2,157 leads were extracted from disclosed contact details of practitioners participating in data warehousing and CRM conferences. In MS Excel, I implemented a database for the leads in order to (i) clean the data, (ii) segment the leads, and (iii) record the status and communications with them. Also in MS Excel, I implemented a database for the institutions in order to record the status and communications with them.

Additionally, I surveyed the market of software programs for massive e-mailing. I selected and purchased e-Campaign 2.93.1 Standard edition. E-mails were sent with a sender's e-mail address at Henley Management College. Only official Henley's envelopes and letterheads were used in regular mail. Communications were

personalized. I manually signed each letter. All the activities related to the correspondence like quality checks (e.g. address in labels vs. name in letter, signature in letter), folding and stamping were performed by my mother.

4.3.11 Pilot Study

The third version of the questionnaire, the TeleForm web-based infrastructure for data collection, the design of the outbound communications, the e-Campaign program and the MS Excel spreadsheets that I created for tracking responses were pretested in a pilot study. Of the 3,517 leads, I identified those individuals with whom I had a professional relationship and I felt comfortable using this past experience to gain further insights if needed. This selection process concluded with 169 leads for the pilot study. The pilot survey started on July 21st, 2003 and ended in September 9th, 2003 (8 weeks).

I received 33 responses, which implied a response rate of 19.53%. The inbound weekly throughput was 4.13 cases. Cases were filtered as stated in the DATA VALIDATION AND PREPARATION section. Cases passing the automatic filters were visually inspected looking for patterns indicating anomalies like all ones and then all twos. There were 25 valid cases for analysis, four of which (16%) were anonymous (see Table 4-34). Only e-mail communications were used.

In relation to the scales, three types of issues were detected in some items: Lack of consistency in scores of some reverse code items, concerns about the wording, and anomalous alphas. Lack of consistency in scores of some reverse code items emerged after transformation by comparing them with the other items in the same summated scale. Concerns about wording were captured in follow up communications with the informants. Cronbach's alpha for this small sample was used just as an indicator of issues in the internal consistency of the items in the summated scale and only in cases of negative value.

		quality	enacenh
N	Valid	25	25
	Missing	0	0
Mean		4.4510	5.8474
Std. Error of Mean		.17083	.26502
Median		4.5667	6.0000
Mode		4.73	6.00
Std. Deviation		.85417	1.32508
Variance		.730	1.756
Skewness		-.118	-.477
Std. Error of Skewness		.464	.464
Kurtosis		-1.119	-.921
Std. Error of Kurtosis		.902	.902
Range		2.73	4.62
Minimum		3.07	3.38
Maximum		5.80	8.00
Sum		111.27	146.19
Percentiles	20	4.4364	3.4533
	40	5.9143	4.3876
	60	6.4909	4.7333
	80	7.0182	5.2874

Table 4-34 Satistics of QUALITY and ENACENH variables in the pilot study

Details on the implemented changes are in APPENDIX D: PILOT RESEARCH STUDY. QUESTIONNAIRE.v3. Overall, I reduced the percentage of items with code reverse from 32.5% to 12.5% leaving the following summated scales without code reverse items: COMP, INT, DKNOWE, PKNOWE, and ICPLXE. Also, I removed demographic items (i.e. position, level, DBMS supporting the data warehouse, firm's industry) not involved in hypothesis testing. Name and organization items were changed to optional because in a few cases I obtained senseless input (i.e. 'GIGO' responses). Qualification and personal items were unified in a single section at the end of the questionnaire. Implemented changes originated a fourth version of the questionnaire. Additional lessons learned were:

- Reminders were absolutely necessary.
- E-mails sent with a 'read notification required' option helped monitoring the progress.

- Monitoring the progress was a daily task at the lead level.
- TeleForm web-application, weekly downloads, and e-Campaign software worked well. No change was required.
- MS-Excel worksheets were very helpful. No changes were required.

The 25 valid cases of the pilot study were not included in the sample of the large-scale survey. The only way of assessing the effectiveness of the introduced changes would have been by conducting parallel surveys with the third and fourth versions of the questionnaire and comparing the results. I did not do it. Nevertheless, in APPENDIX E: SURVEY QUESTIONNAIRE.v4 I have included details on the alphas that I had in the large-scale sample. These alphas correspond to the summated scales in the fourth version. Again, it is not appropriate to compare alphas of a pilot study with 25 cases with the alphas of a sample with 112 cases. However, it is interesting to note that (i) when I did nothing as a result of the pilot study, the alphas in the sample never improved with respect to the reference in the pilot, and (ii) when I introduced changes, the alphas in the sample were never worse with respect to the reference in the pilot. In summary, this pilot test discovered issues in wording, questionnaire lay-out and response format which are the basic goal of a pilot test according to [Bagozzi 1996].

4.3.12 Sample Size

With respect to the targeted sample size, it is obvious that increases in sample size make the results more accurate by reducing the standard error but this is not a simple linear association (e.g. [Curwin and Slater 1996]). So, I needed guidelines in setting my objective for my critical sample size in terms of the number of valid responses (outliers discounted). Given that generalization is a key objective of this study the efficient sample size has to be determined in order to enable inferential assessment. Three decisions had to be made in this sense: degree of confidence DC, level of precision DP, and the amount of variability *s*. [Hair Jr. *et al.* 2003]. I focused on QUALITY and ENACENH as the two main variables in my research model to look at. In the pilot both variables meet the normality requirements (see Table 4-34) and ENACENH has a higher variability (s = 1.32) than QUALITY. (s = .85). In order to be conservative I took ENACENH standard deviation for the calculation of the

ENACENH has a higher variability (s = 1.32) than QUALITY. (s = .85). In order to be conservative I took ENACENH standard deviation for the calculation of the efficient sample size. Because the variables follow a normal distribution, I decided to take a 95 percent confidence level DC of being correct (i.e. 2 standard errors). Finally, I specified 1/3 of a unit for the amount of acceptable error DP between the sample value and the true population value. The sample size SS is calculated as follows

Sample size (SS) = $(DC \times s / DP)^2$

SS = $(2 \times 1.32 / .33)^2 = 64$

Therefore 64 is the minimum sample size in order to meet the specified precision and confidence goals. In addition, the standard practice in doctoral dissertations is to have more than 100 at the low end. I also identified a high end applying a guideline (5 cases per parameter) for structural equation modeling [Hair Jr. *et al.* 1998], SEM in short, which in my case demanded 345.

4.3.13 Survey Administration

This research is cross-sectional in that all the observations for all the variables were taken in a short period of time [Spector 1981]. The administration of the survey started on September 26th, 2003 and ended in May 8th 2004.

I also approached 21 associations and publications. The outcome was that one publication and seven associations supported this research. The DM Review publication supported this research in their web edition. The following seven associations acted as channels supporting this research: DataWarehouse.com (in their web-portal), The Data Management Association (in their web-portal and in their monthly communication to members), The Project Management Institute Information Systems SIG (e-mail to their members), Colegi Oficial d'Engineers en Informatica de Catalunya (e-mail to members), Canadian Information Processing Society (monthly communication to associates), Asociacion de Tecnicos en Informatica (e-mail to members), and The Global Billing Association (monthly communication to members).

In order to increase responses, I implemented a multi-mode strategy to obtain responses from individuals reachable through e-mail (see APPENDIX H: OUTBOUND COMMUNICATIONS vs. VALID CASES CHARTS). This strategy consisted of letterheads by regular mail to the leads that did not respond to the first follow up e-mail. After, sending the mail, I sent follow up e-mails making reference to "my previous letter". This multi-mode approach proved to be successful in the sense that I saw responses coming from these individuals. However, there is no doubt about the most successful strategy that I implemented: On September 26, 2003 the questionnaire was distributed to the 15 individuals attending a meeting, they provided their input before leaving the room, and this resulted in six valid cases. My conclusion is that one way for increasing the inbound weekly throughput to take advantage of having the individuals together in the same place and asking them for fulfilling the questionnaire right there.

The 3,517 collected leads required 8,529 outbound communications with a weekly throughput of 328.24 communications. See Table 4-35 for a description of the outbound communications. A partial assessment of the data quality of the collected leads is given by the number of leads not reached. This was 10.49% in e-mails and 7.14% in regular mail. Although, I do not have standards for comparison, I found them to be at a very satisfactory level. In monetary terms, considering that the average cost per letter was 1.06 USD, this means that the regular mail cost 1,531.00 USD with a waste of 109 USD due to poor data quality.

# Leeds (e-mail or e-mail&postal) in preexisting lists			2157					
# Leeds (e-mail) from own elaboration			1360					
# Leeds e-mail not reached (delivery failure)			369	10.49%				
# Leeds mail not reached (Return to sender)			103	7.14%				
# First contacts (discounted delivery failures)				3045				
# Follow up contacts				5012				
# Outbound communications					8529			
# e-mails		7071						
# mails		1443	(out of	2157	) # "Return to sender		103	7.14%
# Paper forms		15	(out of	1360	)			
# WEEKS	26							
OUTBOUND THROUGHPUT:	328.04							

Table 4-35 Descriptives of outbound communications

There were 161 responses which implied an overall response rate (over first contacts discounted delivery failures) of 4.99%. This is an outstanding achievement compared to the industry standard of 1.87% [DMA 2003]. The inbound weekly throughput was 5.85 cases. Cases were filtered as stated in DATA VALIDATION AND PREPARATION. Cases passing the automatic filters where visually inspected looking for patterns indicating anomalies. There were 112 valid cases for analysis (see Table 4-36 for descriptives). The percentage of anonymous valid cases was surprisingly low (12.5%) considering that identification was optional. I attribute this low rate to a success in the stimulus offering a summary of research findings (83.04%).

# Cases requesting summary	93	83.04%				
# Cases disclosing organization	83	74.11%				
# Case not disclosing org	29	25.89%				
# disclosed organizations	78					
# Cases coming from a single org	72	64.29%				
# Cases w. a common org	11	9.82%	(i.e. They come from.....		6	organizations)
		100.00%				

Table 4-36 Descriptives of the 112 valid cases

I should comment that I had unfounded optimistic expectations about the weekly throughput of inbound communications. To my dismay, the low pace of responses (see APPENDIX I: MULTIMODE-OUTBOUND COMMUNICATIONS) illustrated by the weekly throughput of 5.85 cases with respect to the weekly throughput of 328.24 outbound communications was absolutely unexpected. Consequently, as soon as I exceeded the low end N_L (i.e. N_L>100), I made the following decisions: (i) I stopped outbound communications, (ii) I did not stop inbound cases – in fact, I received 8 additional valid cases, and (iii) I started the analysis of the 112 valid cases.

4.4 SUMMARY OF RESEARCH METHODOLOGY CHAPTER

This chapter has provided a comprehensive description of the planned and executed research steps. The papers that I produced as a by-product and the attendees to the numerous review presentations that I offered give sufficient evidence that the described planned steps were done a-priori as opposed to an outcome from a trial-and-error approach. Obviously, flexibility in the direction is important and I had to

implement changes in the planned steps. In general, the described research design accomplished the overall goal of guiding my research action in an effective way.

This research is fundamentally positivistic and required lengthy qualitative interventions for validation and explanatory purposes. The process of elaboration of the questionnaire and the required artifacts supporting the survey administration were detailed. The sample size was 112 valid cases. The plan for conducting the analysis was also presented, which should be consulted when reading the next chapter.

5. DATA ANALYSIS

5.1 INTRODUCTION

This chapter and referred appendixes cover all the data analysis. The DATA VALIDATION AND PREPARATION section covers the pre-analysis activities. The RELIABILITY AND NORMALITY ANALYSES section includes the identification of outliers. The REPRESENTATIVENESS OF THE SAMPLE DATA section covers analyses and assessments with respect to the population. The PHASES IN THE ANALYSIS section contains the overview of the research plan. This plan had six phases and each phase has a section.

There is a GENERALIZATION OF RESEARCH FINDINGS section where this important aspect is assessed. Finally there is a SUMMARY OF HYPOTHESES NOT SUPPORTED section leaving to the next chapter a comprehensive summary of the research findings.

Anylises were done using the following software applications: SPSS [SPSS 2003] and AMOS [Arbuckle 2003], calculations with SPSS were replicated using SAS/STAT [SAS 2005].

5.2 DATA VALIDATION AND PREPARATION

On a weekly basis I received a MS-Excel spreadsheet by e-mail with all the responses entered thus far. No processing was done to the data before reaching me. The spreadsheet had a row per case (i.e. response) and a column for each item in the questionnaire. Overall, I received 161 responses before I stopped accepting any more. Received cases were moved to another spreadsheet where I implemented the following actions:

- **Automatic data validation**: I implemented a formula in order to filter qualified responses for the unit of analysis of this research (see CONTEXT, SITUATION AND

UNIT OF ANALYSIS RESEARCH SPECIFICATIONS). Invalid cases did not progress to the next steps and did not participate in the analysis.

- **Visual inspection**: I visually inspected cases considered valid after the automatic data validation looking for patterns indicating anomalies (e.g. no input, all ones, duplicates). Forty nine out of the 161 responses were considered invalid and did not progress to the next steps.
- **Time stamping and identification:** I allocated the downloading date to each new received case. On occasions, respondents did not provide their names but they did provide their e-mail addresses. When it was possible to identify their organization or their names from the e-mail address, I did not consider these responses as anonymous.
- **Nonresponses treatment:** For items item_1.v4 to item_54v4 there were 1.91% nonresponses. Nonresponses in Likert scales were addressed with mean substitution (e.g. [Hair Jr. *et al.* 1998]). Nonresponses in 100 constant sum scales where the provided responses reached 100 were addressed with (number) zero substitution. Nonresponses in 100 constant sum scales where the provided responses did not reach 100 were addressed by contacting the respondent, when possible, and kindly asking for the missing values. When this contact was not possible or I did not get response, I addressed these no responses with mean substitution. With this treatment I reduced the 1.91% to 0%.
- **Sample error treatment:** In 100 constant sum scales, error was the difference between 100 and the calculated total. These errors were addressed by contacting the respondent, when possible, and kindly asking for the right values. When this contact was not possible or I did not receive a response, I addressed these errors by evenly distributing the error among the items (e.g. 60&30 became 65&35). This action happened in 1.4% of the items item_49i.v4 to item_54iv.v4 (i.e. 28 items) in the sample N=107 and the error never exceeded the value 10 over 100.
- **Data Transformation**: Reverse code items were automatically transformed in a second MS-Excel spreadsheet. For example, for a 7-point Likert scale the formula was *transformed score = 8 - score.*

5.3 RELIABILITY AND NORMALITY ANALYSES

A major objective of this research is to be able to generalize, with respect to the population, the findings obtained in the sample. A normal distribution in the sample data along with an acceptable sample size and sample representativeness will allow me to generalize the research findings via statistical inference. After analyzing the reliability and normality of the data (step #7 in [Hair Jr. *et al.* 2003] pag. 176), it might be necessary to introduce additional purification in the scales (step #8 in [Hair Jr. *et al.* 2003] pag. 176). This purification is usual in exploratory research (e.g. [Hair Jr. *et al.* 2003]). Furthermore, reliability refers to the *results* obtained with an evaluation instrument and not to the instrument itself. Thus, it is more appropriate to speak of the reliability of 'test scores' or the 'measurement' than of the 'test' or the 'instrument' [Henson, Kogan and Vacha-Haase 2001]. Reliability may vary with different administrations (i.e. it is contingent upon sample characteristics) [Henson *et al.* 2001].

Five valid cases out of the 112 had serious departures from the normal distribution across several scales and were not used in the analysis. Therefore the sample size of this research is N= 107. The purification of scales consisted in the removal of the items with unacceptable item-to-total correlation or inter-items correlation. Thresholds were >5 for the former and >3 for the latter [Hair Jr. *et al.* 2003]. The scales of the first order latent constructs CUST, EQUI, and TRUST were purified in the following way:

CUST: Items item_4.v4, item_5.v4, and item_6.v4 were removed.
EQUI: Item item_19.v4 was removed.
TRUST: Item item_21 was removed.

After purification and removal of outliers, the reliability of the sample data for the quality first order latent constructs (see Table 5-1) was acceptable according to the guidelines for exploratory research (alpha > .60) [Hair Jr. *et al.* 2003]. I did not consider the slightly lower value in INT as concerning. Skewness and kurtosis were the statistics used for assessing the normality of the sample data. Values outside the +/- 1 range for skewness and +/-3 for kurtosis are deemed as departures from

normality [Hair Jr. *et al.* 2003]. Overall, the normality of the sample data for the quality first order latent constructs (seeTable 5-1) was acceptable. I did not consider the slight departures in the skewness in EQUI and RICH as concerning. However, the skewness in TRUST was definitely a concern. This issue was satisfactorily removed after the unidimensionality analysis.

		Alpha >.70 (>.60 for exploratory research) item-to-total correlations > .5 and inter-items correlation > .3				S in +/- 1 and K in +/- 3	
	SCALE	Before purification/outliers		After Purification and Outliers Removal			
		Alpha	# Items	Alpha	# Items	Z Skewness	Z Kurtosis
CUST	Customer Insights	0.47	6	0.60	3	0.0390	-1.3241
COMP	Competitor Insights	0.71	3			-0.1208	-1.0262
ENAC	Customer relationship problem enactment	0.83	4			-0.5794	-0.2126
INT	Data Integration	0.58	3			-0.3448	-0.7373
EQUI	Data Equivocality	0.51	3	0.75	2	1.0148	1.3591
RICH	Data Richness	0.78	5			-1.0436	0.4592
TRUST	Data Trustworthiness	0.55	3	0.73	2	-1.2263	1.0674

Table 5-1 Reliability and normality in quality constructs before unidimensionality analysis

Regarding the cognitive constructs in the research model, no purification was needed. After removal of outliers, the reliability of the sample data for the cognitive first order latent constructs (see Table 5-2) was acceptable according to the guidelines for exploratory research (alpha > .60) [Hair Jr. *et al.* 2003]. I did not consider the slightly lower value in PKNOWE as of concern. In order to increase my level of confidence in the self-efficacy data I evaluated the reliability of the three subscales [Wolf 1997].

Overall, the normality of the sample data for the cognitive first order latent constructs (see Table 5-2) was acceptable. However, the skewness in DKNOWE was definitely a concern. This issue was satisfactorily removed after the unidimensionality analysis.

		Alpha >.70 (>.60 for exploratory research) item-to-total correlations > .5 and inter-items correlation > .3				S in +/- 1 and K in +/- 3	
	SCALE	Before purification/outliers		After Purification and Outliers Removal			
		Alpha	# Items	Alpha	# Items	Z Skewness	Z Kurtosis
DKNOWE	Customer Relationship Problem Declarative Knowledge Enhancement	0.95	3			-1.1491	1.0337
PKNOWE	Customer Relationship Problem Enactment Procedural Knowledge Enhancement	0.59	4			-0.4732	0.0956
ICPLXE	Customer Relationship Problem Integrative Complexity Enhancement	0.82	2			-0.6449	0.0746
LSELFE	Low Customer Relationship Problem Enacting Self-Efficacy	0.82	4			-0.7222	0.1426
MSELFE	Medium Customer Relationship Problem Enacting Self-Efficacy	0.84	4			-0.5601	0.3043
HSELFE	High Customer Relationship Problem Enacting Self-Efficacy	0.80	4			-0.3638	-0.0091
SELFE	Customer Relationship Problem Enacting Self-Efficacy	0.92	12			-0.3695	-0.3634

Table 5-2 Reliability and normality in cognitive constructs before unidimensionality analysis

Regarding the behavioral constructs in the research model, no purification was needed. After removal of outliers, the reliability of the sample data for the behavioral first order latent constructs (see Table 5-3) was acceptable according to the guidelines for exploratory research (alpha > .60) [Hair Jr. *et al.* 2003].

Overall, the normality of the sample data for the behavioral first order latent constructs (seeTable 5-3) was acceptable except in AA. Fortunately, [the departure from] normality in AA was not an issue in most of the types of analysis conducted (cluster and discriminant analysis). In ANOVA analyses this issue is reported.

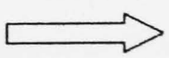
		Alpha >.70 (>.60 for exploratory research) item-to-total correlations > .5 and inter-items correlation > .3				S in +/- 1 and K in +/- 3	
	SCALE	Before purification/outliers		After Purification and Outliers Removal			
		Alpha	# Items	Alpha	# Items	Z Skewness	Z Kurtosis
SCAN/FOCUS	Intensity of data warehouse scanning	0.66	2			0.2253	-0.7993
REP	Intensity of data warehouse representativeness search heuristic	0.66	4			0.4615	0.0887
AVBLE	Intensity of data warehouse availability search heuristic	0.67	4			0.7102	0.7099
AA	Intensity of data warehouse anchoring and adjustment search heuristic	0.78	4			1.4583	3.3169
POSI	Intensity of data warehouse positivy search heuristic	0.64	4			0.9589	1.8234

Table 5-3 Reliability and normality in behavioral constructs before unidimensionality analysis

The variable CRM job function (JOBFUNC) has four groups (sale support, marketing support, customer service support, and other CRM functions). A requirement for ANOVA analyses is that groups should have similar size [Hair Jr. *et al.* 2003] (ratio of 1.5 or higher may be a problem). Therefore, in order to satisfy the equality in size requirement, I randomly withdrew a sample of 50% (see Figure 5.1) from the larger group (i.e. marketing support)

Between-Subjects Factors

		Value Label	N
61 What is the primary CRM function of your personal job?	1	Sales Support	21
	2	Marketing Support	49
	3	Customer Service Support	21
	4	Other CRM Functions	16



Between-Subjects Factors

		Value Label	N
61 What is the primary CRM function of your personal job?	1	Sales Support	21
	2	Marketing Support	24
	3	Customer Service Support	21
	4	Other CRM Functions	16

Figure 5.1Even size in JOBFUNC groups

The variable number of supported CRM data warehouse functions (DWFUNC) has a multiple answer format. I considered that the most interesting groups were the group that I labeled as "just one CRM function supported by the data warehouse" and the group that I labeled as "more than one CRM function supported by the data warehouse". I implemented a dummy variable [Hair Jr. *et al.* 2003], labeled dummy60, in order to create these two groups. In order to satisfy the equality in size requirement, I randomly selected 12 cases in the larger group (see Figure 5.2).

Between-Subjects Factors

		Value Label	N
60 DUMMY Which CRM functions are supported by your data warehouse?	1	Just one CRM function	12
	2	More than one CRM Function	95

Between-Subjects Factors

		Value Label	N
60 DUMMY Which CRM functions are supported by your data warehouse?	1	Just one CRM function	12
	2	More than one CRM Function	12

Figure 5.2 Even size in DWFUNC groups

At this stage, the sample data appeared to follow a normal distribution (with the few exceptions of TRUST, DKNOWE, and AA) and seemed reliable.

5.4 REPRESENTATIVENESS OF THE SAMPLE DATA

The unit of analysis of this research defines the profile of the members of the population from which a sample was extracted. It seems reasonable to assess the representativeness of this sample by looking at the attributes of this sample and comparing them with the definition of the unit of analysis of this research. A few descriptive analyses profiling the sample and an assessment of their representativenes value follows:

. **The profile of the sample respondents is consistent with that of the study's defined population in terms of geographical area.**

The population has a worldwide presence as indicated in reports about the spending in data warehouses (e.g. [Soejarto 2002]). The market size in 2001 had the following distribution per area: NA North America 52.8% Europe 33.1% Asia-Pacific 5.4% Japan 4.9% LA Latin America 1.8% ROW Rest of the World 2.0% [Soejarto 2002]. In the sample, I was able to trace geographical area in 86 cases (see Figure 5.3).

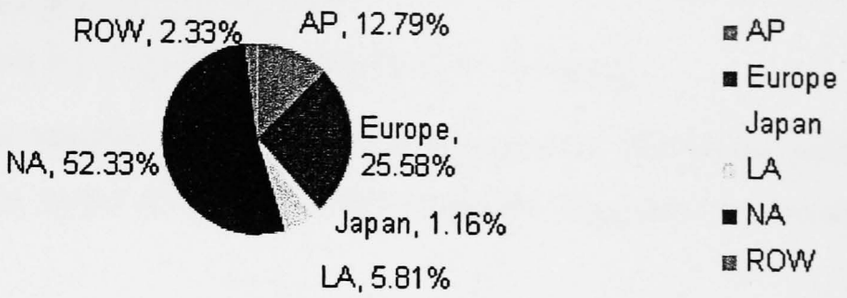


Figure 5.3 Geographical areas of informants in sample

The referred report with the 2001 market size in terms of spending can be considered as a reference in order to assess the representativenes of the sample in terms of geographical diversity. Interestingly, the proportions are very similar. Therefore, it seems reasonable to conclude that the sample is representative of the population in terms of geographical areas and cases per area. This geographical dispersion came from twenty one countries (see Table 5-4).

# Disclosed countries in valid cases	21					
# Valid cases disclosing country	86	80.37%				
# Valid cases not disclosing country	21	19.63%				
# Valid cases coming from a single country	10	11.63%				
# Valid cases w. a common country	76	88.37%	(i.e. They come from.....		11	countries)

Table 5-4 Countries of informants in sample

- **The sample is cross-sectional in terms of data warehouses.**

Because the informants in the population work for organizations that own data warehouses, the more different organizations in the sample the better cross-sectional representation. The ideal would be to have one respondent per organization. In this research, the sample with N=107 comes from at least 69 organizations (see Table 5-5). Not making any assumption about the valid cases that did not disclose their organizations, the 64.49% seems to be a good cross-sectional representation of data warehouses. Furthermore, twelve cases came from seven organizations (i.e. roughly 1.7 cases per organization) increasing the percentage of data warehouses with less than 1.1 respondents per organization to 75.70%.

# Valid cases disclosing organization	81	75.70%				
# Valid cases not disclosing org	26	24.30%				
# Valid cases coming from a single org	69	64.49%				
# Valid cases w. a common org	12	11.21%	(i.e. They come from....	7	organizations)	
#Valid cases	107	100.00%				
# Disclosed organizations in valid cases	76					

Table 5-5 Organizations of informants in sample

. **The profile of the sample respondents is consistent with that of the study's defined population in terms of types of CRM functions supported by the data warehouse.**

Market analysts (e.g. [Wardley and Blumstein 2004, Maoz, DeSisto, Marcus, Herschel, Kolsky, Thompson and Berg 2003, Graham, Latimer, Biscotti, Correia, Eschinger, Pang and Topolinski 2004]) have a rare consensus in the types of CRM functions. The classical three major functions are sales support, marketing support, and customer support. The frequencies per CRM function supported by data warehouse are in Table 5-6. I am unaware of information about the relative proportions in the population among such CRM functions in terms of number of data warehouses supporting them.

60 DUMMY Which CRM functions are supported by your data warehouse?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Just one CRM function	12	11.2	11.2	11.2
	More than one CRM Function	95	88.8	88.8	100.0
	Total	107	100.0	100.0	

		60 DUMMY Which CRM functions are supported by your data warehouse?	
		Just one CRM function	More than one CRM Function
		Count	Count
60_i Which CRM functions are supported by your data warehouse? Sales support	0	10	8
	Sales Support	2	87
60_ii Which CRM functions are supported by your data warehouse? Marketing support	0	9	8
	Marketing Support	3	87
60_iii Which CRM functions are supported by your data warehouse? Customer Service support	0	6	32
	Customer Service Support	6	63
60_iv Which CRM functions are supported by your data warehouse? Other CRM functions	0	11	48
	Other CRM functions	1	47

Table 5-6 Types of CRM functions supported by the data warehouses

Because the data warehouses in the sample support the three major CRM functions in the population, it seems reasonable to conclude that the sample is representative of the population in terms of types of CRM functions supported by the data warehouses.

. **The profile of the sample respondents is consistent with that of the study's defined population in terms of types of CRM functions performed by the informants.**

I am unaware of information about the relative proportions in the population among CRM functions in terms of number of practitioners performing them. The frequencies per CRM function performed by the informants are given in Table 5-7.

61 What is the primary CRM function of your personal job?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Sales Support	21	19.6	19.6	19.6
	Marketing Support	49	45.8	45.8	65.4
	Customer Service Support	21	19.6	19.6	85.0
	Other CRM Functions	16	15.0	15.0	100.0
	Total	107	100.0	100.0	

Table 5-7 Types of CRM functions performed by informants

Because the informants in the sample perform the three major CRM functions in the population, it seems reasonable to conclude that the sample is representative of the population in terms of types of CRM functions performed by the informants.

. **All the respondents in the sample are not novice practitioners in the context of this research.**

The subjective judgments of informants should be founded in subject matter expertise in CRM supported by a data warehouse, which is the context of this research (e.g. [Kolb 1984, Sticht 1976, Bandura 1986]). The data warehouse and CRM experiential profile of informants is in Table 5-8.

The experiential profile of informants in Table 5-8 shows that they are not novices performing CRM tasks supported by a data warehouse and are therefore qualified people for providing input.

. **All the respondents in the sample perform activities in the situation-within-context of this research.**

The subjective judgments of informants should be founded in the informants' experience enacting CRM problems and in the informants' information search

behavior searching/requesting customer information from the firm's data warehouse for enacting CRM problems, which is the situation-within-context in this research (e.g. [Weick 1993]). The CRM problem enactment and behavioral experiential profile of informants is given in Table 5-9. Data in Table 5-9 is expressed in terms of percentages.

55 How long have you been involved in customer relationship management tasks?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Less Than 6 months	1	.9	.9	.9
	Between 6 months and 1 year	7	6.5	6.5	7.5
	Between 1 and 3 years	21	19.6	19.6	27.1
	Between 3 and 5 years	29	27.1	27.1	54.2
	More than 5 years	49	45.8	45.8	100.0
	Total	107	100.0	100.0	

56. How long have you been using your company's data warehouse customer relationship data to help you understand issues?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Less Than 6 months	3	2.8	2.8	2.8
	Between 6 months and 1 year	20	18.7	18.7	21.5
	Between 1 and 3 years	28	26.2	26.2	47.7
	Between 3 and 5 years	24	22.4	22.4	70.1
	More than 5 years	32	29.9	29.9	100.0
	Total	107	100.0	100.0	

Table 5-8 Data warehouse and CRM experiential profile of informants

	61 What is the primary CRM function of your personal job?				
	Sales Support	Marketing Support	Customer Service Support	Other CRM Functions	Total
	Mean	Mean	Mean	Mean	Mean
59. In a typical week how much time do you spend searching/requesting customer information from your firm's data warehouse?	39	40	45	41	41
58. In a typical week how much time do you spend understanding CRM issues (e.g. challenges, problems)	40	45	50	50	46

Table 5-9 CRM problem enactment and behavioral experiential profile of informants

The experiential profile of informants in Table 5-9 shows that they are engaged in activities enacting CRM problems and searching/requesting customer information from the firm's data warehouse for enacting CRM problems and are therefore qualified people for providing input.

Overall, the sample has a profile that suggests it is a good representative of the population.

5.5 PHASES IN THE ANALYSIS

The research model in Figure 3.1 has the particularity that all the constructs are dependent and independent variables which will require bidirectional analysis in every association. As stated in the RESEARCH STRATEGY section, this research is exploratory since it involves (i) known high order latent constructs with little knowledge of their internal structures, and (ii) the adaptation of existing scales in order to measure such latent constructs. It is confirmatory since it involves theoretically known associations. Rather than a strict dichotomy of exploratory vs. confirmatory, this research can be thought of as an ordered progression [Anderson and Gerbing 1988] where the research model can be viewed as the space of the hypotheses in RESEARCH HYPOTHESES that I wanted to test. The hypotheses testing was planned and executed following six differentiated and sequential phases (see Table 5-10).

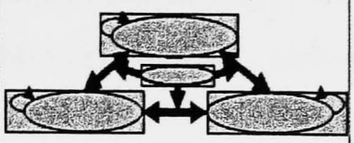
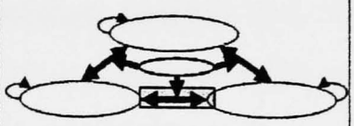
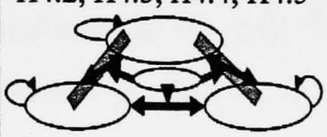
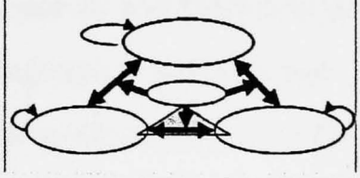
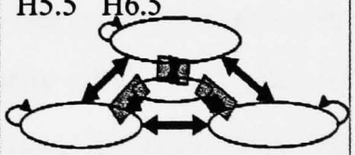
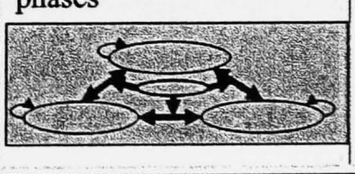
ANALYSIS PHASE	HYPOTHESES	TYPE OF RESEARCH	ANALYSIS TECHNIQUES
I	H1, H2, H3.1, H3.2 	Exploratory	- Pearson's correlations and multivariate regression analysis for multicollinearity issues - Factor and Cluster analysis
II	H4.1 	Confirmatory	- Pearson's correlations and multivariate regression - CFA using Structural Equation Modeling
III	H4.2, H4.3, H4.4, H4.5 	Confirmatory	- Discriminant analysis
IV	H5.1 H6.1 	Exploratory	- ANOVA analysis
V	H5.2 H6.2 H5.3 H6.3 H5.4 H6.4 H5.5 H6.5 	Exploratory	- ANOVA analysis
VI	All supported hypotheses in previous phases 	Confirmatory and Explanatory	- Qualitative analysis

Table 5-10 Six phases in the analysis

5.6 PHASE I: EXPLORATORY ANALYSIS OF INNER CONTRUCT ASSOCIATIONS

5.6.1 Exploratory Analysis of Environmental Constructs

The first exploratory analysis consisted of discovering correlations among the indicators of QUALITY. The expectation was that if there is a structure as hypothesized in hypothesis H1, then a good signal should be that there are correlations among the indicators CUST, COMP, INT, RICH, ENAC, and EQUI. Pearson's correlation analysis (see Table 5-11) showed statistically significant (p <.000 and p <.05) correlations either of a moderate intensity (i.e. R in the range .41 to .70) or a small but definite association (i.e. R is the range .21 to .40). The guidelines in Hair Jr. *et al* [2003] were applied in assessing the intensity of the correlation coefficient R and colored for easy identification. However, a high correlation among independent variables might indicate a multicollinearity issue [Hair Jr. *et al.* 1998].

The next step was to test the intradependencies among the indicators of QUALITY. Multivariate regression analysis was conducted for that purpose. Here the relevant finding was the confirmation of multicollinearity issues (R > .71 p < .000) in ENAC data and RICH data (see Table 5-12). I addressed this disturbing finding by removing the entire ENAC scale and items item_23.v4 and item_27.v4 in the RICH scale.

After clearing the multicollinearity issues, the next exploratory step was to test the unidimensionality of the indicators of QUALITY CUST, COMP, INT, RICH, and EQUI. Exploratory factor analysis was applied to the measurement variables of such indicators. Item_16.v4 in the INT scale was removed due to its null contribution to any factor. After removing this item, loadings on items in first order quality latent constructs show unidimensionality (see Table 5-13). They rank between .71 and .90 which is considered high and a few fall in the rank .41 to .70 which is considered moderate [Hair Jr. *et al.* 2003]. The R^2 variance extracted by this factor solution was 69.646 which exceeds the commonly used .50 threshold value for acceptable reliability [Hair Jr. *et a*l. 1998].

			21- 40 Small but definite	41-70 Moderate	71-90 High	91-100 Very Strong	**Multicollinearity cut-off values for acceptable levels: VIF < 5 Tolerance > .1 R > .7**											
		Quality of DWH CR Data for problem Enactment																
	Pearson's Statistically Significant Correlations	**CUST** (Customer insights)	**COMP** (Competitor insights)			**ENAC** (Customer relationship problem enactment)			**INT** Integration			**EQUI** Equivocality			**RICH** Richness			**TRUST** Trustwothiness		
			R	**Sig.**	**R^2**	**R**	**Sig.**	**R^2**	**R**	**Sig.**	**R^2**	**R**	**Sig.**	**R^2**	**R**	**Sig.**	**R^2**	**R**	**Sig.**	**R^2**
Quality of DWH CR Data for problem Enactment	**CUST (Customer insights)**		0.484	0.000	0.234	0.507	0.000	0.257	0.349	0.000	0.122	-0.230	0.017	0.053	0.352	0.000	0.124	0.240	0.013	0.057
	COMP (Competitor insights)					0.488	0.000	0.238	0.286	0.003	0.082	-0.335	0.000	0.112	0.345	0.000	0.119	0.303	0.002	0.092
	ENAC (Customer relationship problem enactment)								0.614	0.000	0.377	-0.417	0.000	0.174	-0.702	0.000	0.493	-0.471	0.000	0.222
	INT Integration											-0.427	0.000	0.183	0.639	0.000	0.408	0.561	0.000	0.357
	EQUI Equivocality														-0.407	0.000	0.165	-0.464	0.000	0.216
	RICH Richness																	0.625	0.000	0.391
	TRUST Trustwothiness																			

Multicollinearity issues w. ENAC

Table 5-11 Correlations among the indicators of QUALITY

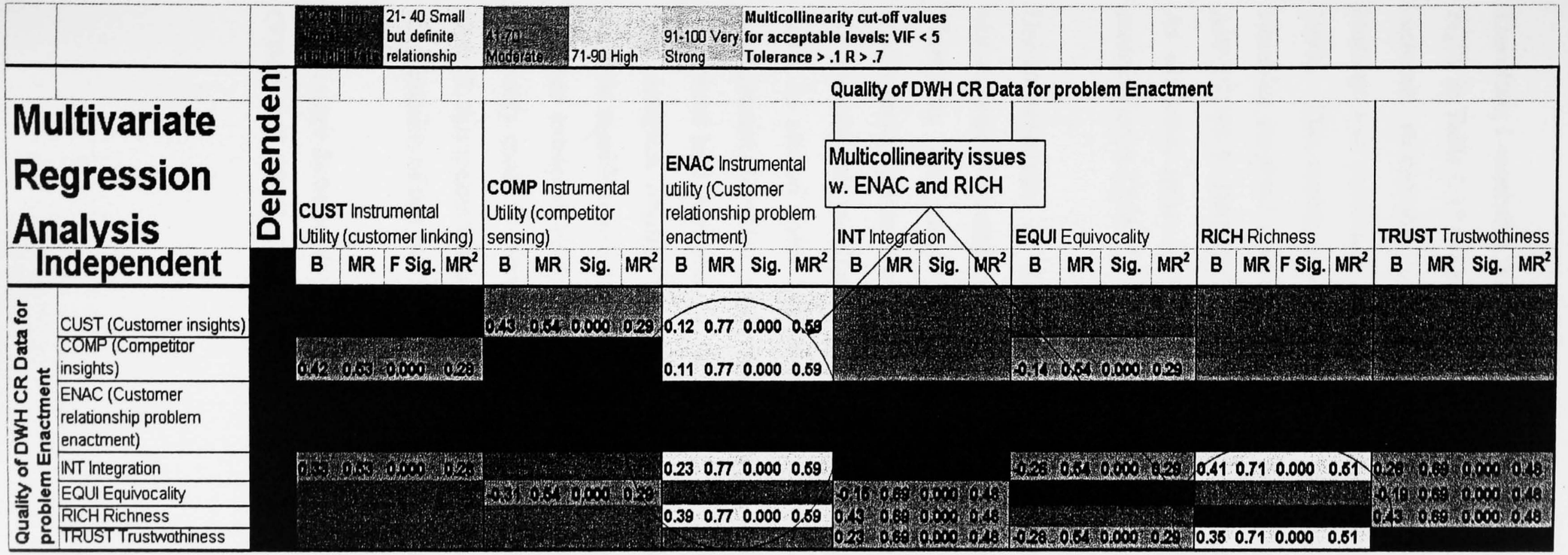


Table 5-12 Mutivariate regression analysis among the indicators of QUALITY

Therefore, I concluded that the five first order scales CUST, COMP, INT, RICH, and EQUI in Table 5-13 is a reliable five factor solution for the quality on data warehouse customer relationship data for problem enactment (QUALITY) construct. Only loadings with practical significance (i.e. exceeding .5) are represented [Hair Jr. *et al.* 1998]. This exploratory factor analysis produced a new operationalization of the variables involved where the TRUST scale was diluted in the other scales (see APPENDIX F: DETAILS ON THE OPERATIONALIZATION OF VARIABLES). As expected, items in the EQUI scale load negatively indicating their negative contribution to QUALITY.

Unidimensionality of the indicators of QUALITY does not address the inner associations advocated in hypothesis H1. Hypothesis H1 advocates two dimensions. Therefore, the measurement variables in Table 5-13 should also show unidimensionality in a two factor solution. Exploratory factor analysis was applied to the measurement variables of QUALITY imposing a two factor solution. Loadings on items in second order quality latent constructs TASKUT and ENACUT show unidimensionality (see Table 5-14). They rank between .71 and .90 which is considered high, between .41 to .70 which is considered moderate, and two of them have marginal loadings (bellow .41) [Hair Jr. *et al.* 2003]. Only loadings with practical significance (i.e. exceeding .5) are represented [Hair Jr. *et al.* 1998]. The R^2 variance extracted by this two factor solution was 47.800 which falls below the commonly used .50 threshold value for acceptable reliability [Hair Jr. *et al.* 1998]. Overall, this means that before claiming support for hypothesis H1, further analysis of the reliability of this two factor solution is required.

Exploratory factor analysis was applied to the first order quality constructs CUST, COMP, INT, RICH, and EQUI.

	Component				
	CUST	COMP	INT	RICH	EQUI
1. It includes customer's motivational data (e.g.needs)	.888				
2. it includes customer's attitudinal data (e.g. trust, satisfaction)	.833				
7. Data about competitors is included (e.g., products, campaigns, channels)		.741			
8. it is a source of data about opportunities for competitive advantage		.777			
9. Allows to respond rapidly to competitors' actions		.691			
3. It does not include relationship measures (e.g. attrition risk, life-time value)			.889		
24. provides new/unanticipated insights for understanding customer relationship problems			.512		
15. allows a 360° view of a customer			.510		
20. can be summarized at different levels				.782	
18. includes accurate data				.758	
22. is organized in a meaningful way.				.747	
25. is a trusted source of customer data				.580	
14. allows you to know the number of customers in your data warehouse				.655	
17. can be interpreted in several ways that lead to different but acceptable customer relationship problem formulations					-.840
26. can support more than one plausible formulation for the customer relationship problems you face					-.839

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 6 iterations

Table 5-13 Rotated five components matrix of items (first order quality constructs unidimensionality)

A two factor solution emerged with loadings on second order latent constructs TASKUT and ENACUT showing unidimensionality (see Table 5-15). They rank between .71 and .90 which is considered high except that INT falls in the rank .41 to .70 which is considered moderate [Hair Jr. *et al.* 2003]. Only loadings with practical significance (i.e. exceeding .5) are represented [Hair Jr. *et al.* 1998]. The R^2 variance extracted by this factor solution was 67.741 which exceeds the commonly used .50 threshold value for acceptable reliability [Hair Jr. *et al.* 1998].

	Component	
	TASKUT	ENACUT
2. it includes customer's attitudinal data (e.g. trust, satisfaction)	.765	
1. It includes customer's motivational data (e.g.needs)	.726	
7. Data about competitors is included (e.g., products, campaigns, channels)	.706	
8. it is a source of data about opportunities for competitive advantage	.658	
9. Allows to respond rapidly to competitors' actions	.618	
20. can be summarized at different levels		.758
25. is a trusted source of customer data		.748
18. includes accurate data		.716
24. provides new/unanticipated insights for understanding customer relationship problems		.673
14. allows you to know the number of customers in your data warehouse		.668
22. is organized in a meaningful way.		.638
26. can support more than one plausible formulation for the customer relationship problems you face		-.585
17. can be interpreted in several ways that lead to different but acceptable customer relationship problem formulations		-.442
15. allows a 360° view of a customer	.557	
3. It does not include relationship measures (e.g. attrition risk, life-time value)	.332	

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. a Rotation converged in 3 iterations

Table 5-14 Rotated two components matrix of measurement variables of QUALITY (second order constructs undimensionality)

Therefore, I concluded that CUST and COMP are indicators of TASKUT and that INT, RICH and EQUI are indicators of ENACUT. Therefore, I feel confident

claiming that TASKUT and ENACUT are dimensions of QUALITY (i.e. H1 is supported).

	Component	
	TASKUT	ENACUT
Customer Insights	.874	
Competitor Insights	.777	
Data Richness		.848
Data Equivocality		-.767
Data Integration		.625

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 3 iterations.

Table 5-15 Rotated two components matrix of first order latent constructs (second order constructs unidimensionality)

As mentioned, after these unidimensionality analyses, a new operationalization of variables was introduced. Therefore, another reliability and normality analyses were conducted. The reliability of the sample data for the first order quality latent constructs CUST, COMP, INT, RICH, and EQUI, second order quality latent constructs TASKUT and ENACUT, and third order quality latent construct QUALITY (see Table 5-16) was acceptable according to the guidelines for exploratory research (alpha > .60) [Hair Jr. *et al.* 2003]. In addition to the Cronbach's alpha, I computed two reliability measures when possible: The construct reliability and the variance extracted [Hair Jr. *et al.* 1998]. The suggested levels considered acceptable are >.7 for the former and >.5 for the latter [Hair Jr. *et al.* 1998]. In all the cases when it was possible to compute these additional measures, the reliability of the data exceeded the cut-off of acceptability. Overall, the normality of the sample data for the quality first order, second order and third order latent constructs (see Table 5-16) was acceptable. I did not consider the slight departures in the skewness in EQUI and RICH as of concern.

		Alpha >.70 (>.60 for exploratory research) item-to-total correlations > .5 and inter-items correlation > .3			Variance extracted > .5	S in +/- 1 and K in +/- 3	
	SCALE	After Factor Analysis					
		Alpha	Composite Construct Reliability	# Items	Variance extracted	Z Skewness	Z Kurtosis
CUST	Customer Insights	0.82	NA	2	NA	0.0390	-1.3241
COMP	Competitor Insights	0.71	NA	3	NA	-0.1208	-1.0262
INT	Data Integration	0.64	NA	3	NA	-0.3448	-0.7373
EQUI	Data Equivocality	0.75	NA	2	NA	1.0148	1.3591
RICH	Data Richness	0.80	NA	5	NA	-1.0436	0.4592
TASKUT	Data Task Utility	0.78	0.81	5	0.68	0.0371	-0.9585
ENACUT	Data Enactment Utility	0.60	0.79	10	0.57	-0.4288	-0.1444
QUALITY	The Quality of DW CR Data for PE	0.74	NA	15	0.70	-0.0815	-0.6632

Table 5-16 Reliability and normality in the indicators of QUALITY after unidimensionality analysis

Distribution statistics for the first order, second order and third order QUALITY latent constructs are in Table 5-17.

	Customer Insights	Competitor Insights	Data Integration	Data Richness	Data Equivocality	Data Task Utility	Data Enactment Utility	Quality of DW Customer Relationship Data for Problem Enactment
N Valid	107	107	107	107	107	107	107	107
Missing	0	0	0	0	0	0	0	0
Mean	3.7150	4.1335	4.4939	5.9022	2.6449	3.9661	4.8283	4.5409
Std. Error of Mean	.18979	.15582	.14236	.08922	.11877	.14509	.06746	.07703
Median	3.5000	4.3333	4.6667	6.2000	2.5000	3.8000	4.8000	4.6000
Std. Deviation	1.96324	1.61179	1.47256	.92293	1.22861	1.50082	.69779	.79681
Variance	3.854	2.598	2.168	.852	1.509	2.252	.487	.635

Table 5-17 Distribution statistics on the indicators of QUALITY

5.6.2 Exploratory Analysis of Cognitive Constructs

The second exploratory analysis consisted of discovering correlations among the indicators of ENACENH. The expectation was that if there is a structure as hypothesized in hypothesis H2 then a good signal should be that there are correlations among the indicators DKNOWE, PKNOWE, ICPLXE, and SELFE. Pearson's correlation analysis (see Table 5-18) showed statistically significant (p <.000) correlations of a moderate intensity (i.e. R in the range .41 to .70). The guidelines in [Hair Jr. *et al.* 2003] were applied in assessing the intensity of the correlation coefficient R and colored for easy identification.

		0-.20 Small, almost negligible/Sta	21- 40 Small but definite relationship	41-70 Moderate	71-90 High	91-100 Very Strong	Multicollinearity for acceptable le Tolerance > .1 R

	Pearson's Statistically Significant Correlations	CRP Sense-making Enhancements											
		DKNOWE CRP Declarative Knowledge Enhancement			**PKNOWE** CRP Enactment Procedural Knowledge Enhancement			**ICPLXE** CR Problem Integrative Complexity Enhancement			**SELF-E** CR Problem Enacting Self-Efficacy		
		R	Sig.	R^2	R	Sig.	R^2	R	Sig.	R^2	R	Sig.	R^2
CRP Sense-making Enhancements	DKNOWE CRP Declarative Knowledge Enhancement				0.631	0.000	0.399	0.562	0.000	0.316	0.572	0.000	0.327
	PKNOWE CRP Enactment Procedural Knowledge							0.607	0.000	0.368	0.650	0.000	0.423
	Integrative Complexity Enhancement										0.617	0.000	0.381
	SELF-E CR Problem Enacting Self-Efficacy												

Table 5-18 Correlations among indicators of ENACENH

The next step was to test the intradependencies among the indicators of ENACENH. Multivariate regression analysis was conducted for that purpose. Here the relevant finding was the discovery of multicollinearity issues (R > .71 p < .000) in PKNOWE data (see Table 5-19). I addressed this disturbing finding by removing item_31.v4 in the PKNOWE scale.

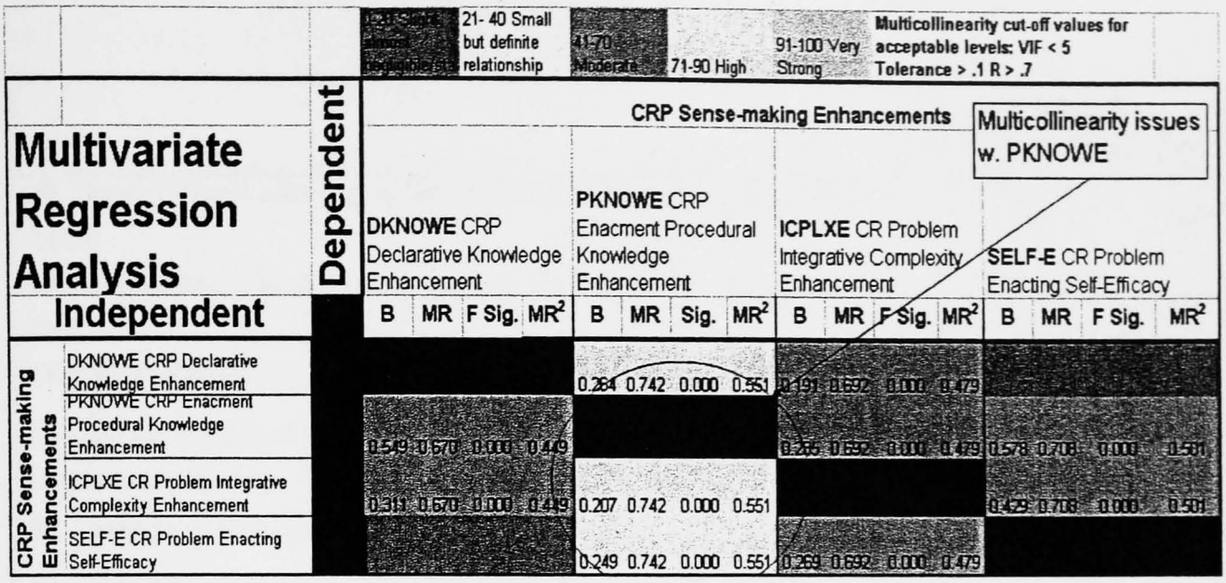


Table 5-19 Mutivariate regression analysis among indicators of ENACENH

After clearing the multicollinearity issue, the next exploratory step was to test the unidimensionality of the indicators of ENACENH DKNOWE, PKNOWE, ICPLXE, and SELFE. Exploratory factor analysis was applied to the measurement variables of such indicators. Item_33.v4 in the PKNOWE scale was removed due to its null contribution to any factor. After removing this item, loadings show unidimensionality in a two factor solution (see Table 5-20). All the items in the SELFE felt in one dimension while the items of DKNOWE, PKONWE, and ICPLXE felt in the second dimension. Only loadings with practical significance (i.e. exceeding .5) are represented [Hair Jr. *et al.* 1998]. The R^2 variance extracted by this factor solution was .58376 which exceeds the commonly used .50 threshold value for acceptable reliability [Hair Jr. *et al.* 1998]. Therefore, I concluded that the first order scales KNOWENH and SELFE in Table 5-20 is a reliable two factor solution for the customer relationship problem sense making enhancements (ENACENH) construct which gives support for hypothesis H2.

	Component	
	SELFE	KNOWENH
28. I have a clearer sense of customer relationship problems (i.e., better focus).		.888
30. I have better insights into the customer relationship threats facing my firm		.886
29. I reach sharper final interpretations of customer relationship problems		.871
34. I am more effective focusing on the key aspects of customer relationship problems		.630
32. I analyze more a customer relationship problem		.620
35. My descriptions of customer relationship problems include more factors		.604
36. My descriptions of customer relationship problems include more interconnections among the factors		.571
43. I can identify new questions whose answers might give new insights	.765	
42. I can reach a plausible customer relationship problem statement under moderate time pressure	.721	
46. I can develop a plausible statement about a high stakes (e.g., risk of losing a market segment) customer relationship problem	.704	
44. I can explain cause-effect relationships in a problem when I have limited information	.702	
40. I can confirm my preliminary conclusions for common problems	.686	
41. I can identify new factors contributing to a customer relationship problem	.683	
45. I can understand radically new types of customer relationship problems	.674	
38. I can formulate a plausible customer relationship problem statement when I have the data that I need	.669	
39. I can determine if there is available data to answer common questions	.657	
47. I can explore different patterns or trends in customer relationship data	.643	
48. I can analyze customer relationship data in a situation where there are competing goals/objectives about the desired customer relationship status (e.g., keep or terminate customer relationship)	.598	
37. I can recognize common factors contributing to a customer relationship problem	.585	

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 3 iterations.

Table 5-20 Rotated two components matrix of measurement variables of ENACENH (first order cognitive constructs unidimensionality)

As a consequence of these findings, a new operationalization of the variables involved was produced where the DKNOWE, PKNOWE, and ICPLXE scales were merged in the KNOWENH scale (see APPENDIX F: DETAILS ON THE OPERATIONALIZATION OF VARIABLES).

As mentioned, after these unidimensionality analyses, a new operationalization of variables was introduced. Therefore, another set of reliability and normality analyses were conducted. The reliability of the sample data for the indicators of ENACENH KNOWENH, and SELFE (see Table 5-21) was acceptable according to the guidelines for exploratory research (alpha > .60) [Hair Jr. *et al.* 2003]. In addition to the Cronbach's alpha, I computed two reliability measures when possible: The construct reliability and the variance extracted [Hair Jr. *et al.* 1998]. The construct reliability measure was not possible to compute. The suggested level considered acceptable for the variance extracted is >.5 [Hair Jr. *et al.* 1998]. In the only case when it was possible to compute this additional measure, the reliability of the data exceeded the cut-off of acceptability.

Overall, the normality of the sample data for the quality first order, second order and third order latent constructs (see Table 5-21) was acceptable.

		Alpha >.70 (>.60 for exploratory research) item-to-total correlations > .5 and inter-items correlation > .3			Variance extracted > .5	S in +/- 1 and K in +/- 3	
	SCALE	After Factor Analysis					
		Alpha	Composite Construct Reliability	# Items	Variance extracted	Z Skewness	Z Kurtosis
KNOWENH	Customer Relationship Problem Enactment Knowledge Enhancement	0.90	NA	7	NA	-0.7846	0.2283
LSELFE	Low Customer Relationship Problem Enacting Self-Efficacy	0.82	NA	4	NA	-0.7222	0.1426
MSELFE	Medium Customer Relationship Problem Enacting Self-Efficacy	0.84	NA	4	NA	-0.5601	0.3043
HSELFE	High Customer Relationship Problem Enacting Self-Efficacy	0.80	NA	4	NA	-0.3638	-0.0091
SELFE	Customer Relationship Problem Enacting Self-Efficacy	0.92	NA	12	NA	-0.3695	-0.3634
ENACENH	Customer Relationship Problem Sense-Making Enhancements	0.94	NA	19	0.58376	-0.4294	-0.3026

Table 5-21 Reliability and normality in indicators of ENACENH after unidimensionality analysis

Distribution statistics for the indicators of ENACENH are given in Table 5-22.

	Customer Relationship Problem Enactment Knowledge Enhancement	Customer Relationship Problem Enacting Self-Efficacy	Customer Relationship Problem Sense-Making Enhancements
N Valid	107	107	107
Missing	0	0	0
Mean	5.3322	6.9801	6.3730
Std. Error of Mean	.10097	.13665	.11524
Median	5.4286	7.0833	6.4211
Mode	6.00	8.00	7.05
Std. Deviation	1.04443	1.41350	1.19209
Variance	1.091	1.998	1.421

Table 5-22 Distribution statistics on the indicators of ENACENH

5.6.3 Exploratory Analysis of Behavioral Constructs

The third exploratory analysis consisted of discovering different behavioral patterns in terms of customer relationship problem sense making information search modes and customer relationship problem sense making information search heuristics.

Regarding mode orientations MODE, cluster analysis was applied to SCAN/FOCUS. Formally, SCAN and FOCUS were two redundant variables as one is the complementary of the other. I arbitrarily decided to work over SCAN. One-way ANOVA analysis provided a statistically significant (p <.000) two cluster solution (see Table 5-23). I labeled cluster #1, 75 cases (70.1%), as FOCUSMO mode orientation due to its low mean, x = 31.52, in SCAN (i.e. high mean in FOCUS). I labeled cluster #2, 32 cases (29.9%), as SCANMO mode orientation due to its high mean, x = 67.67, in SCAN (i.e. low mean in FOCUS). Therefore, I concluded that there are different behavioral patterns MODE in terms of customer relationship problem sense making information search mode orientation FOCUSMO and SCANMO which supports hypothesis H3.1.

	N	Mean	Std. Deviation	Std. Error
FOCUSMO	75	31.5245	13.04710	1.50655
SCANMO	32	67.6719	9.49415	1.67834
Total	107	42.3350	20.53499	1.98519

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	29307.600	1	29307.600	199.940	.000
Within Groups	15391.094	105	146.582		
Total	44698.694	106			

Table 5-23 Cluster analysis of customer relationship problem sense making information search mode orientation

Regarding heuristics, cluster analysis was applied to REP, AVBLE, AA, and POSI. Although the intensity of the data warehouse anchoring and adjustment search heuristic (AA) had departures from normality, this has no effect on cluster analysis [Hair Jr. *et al.* 1998]. One-way ANOVA analysis provided a statistically significant

(p <.000 and p <.001) two cluster solution (see Table 5-24). I labeled cluster #1, 90 cases (84.1%), as TEMPLATE heuristic orientation due to its higher means in REP and AVBLE, x = 25.46 and x = 23.94 respectively, and lower means in AA and POSI, x = 18.62 and x = 17.21 respectively. I labeled cluster #2, 17 cases (15.9%), as TRIAL-and-ERROR heuristic orientation due to its higher means in AA and POSI, x = 42.29 and x = 32.13 respectively, and lower means in REP and AA, x = 15.05 and x = 13.45 respectively. Both representativeness and availability are information search heuristics that have in common the involvement of an information template either as the reference while searching representative information in the former or as the goal of the search in the latter. Anchor and adjustment and positivy are information search heuristics that have in common the trial-and-error method [Newell and Simon 1972]. Therefore, I concluded that there are different behavioral patterns HEUR in terms of customer relationship problem sense making information search heuristic orientation TEMPLATE and TRIAL-and-ERROR which support hypothesis H3.2.

5.7 PHASE II: CONFIRMATORY ANALYSIS OF ASSOCIATION BETWEEN ENVIROMENTAL AND COGNITIVE CONSTRUCTS

The expectation is that there is an association between customer relationship problem sense making enhancements ENACENH and quality on data warehouse customer relationship data for problem enactment QUALITY (i.e. H4.1). Pearson's correlation analysis (see Table 5-25) showed statistically significant correlation of a moderate intensity (i.e. R in the range .41 to .70) between QUALITY and ENACENH with R=.53 (p <.001) which confirms that there is an association between both variables.

		N	Mean	Std. Deviation	Std. Error
REP	TEMPLATE	90	25.4615	11.49193	1.21136
	TRIAL-and-ERROR	17	15.0588	8.05379	1.95333
	Total	107	23.8088	11.63070	1.12438
AVBLE	TEMPLATE	90	23.9478	10.67599	1.12535
	TRIAL-and-ERROR	17	13.4559	5.21434	1.26466
	Total	107	22.2809	10.70753	1.03514
AA	TEMPLATE	90	18.6215	7.78205	.82030
	TRIAL-and-ERROR	17	42.2941	13.65056	3.31075
	Total	107	22.3826	12.43260	1.20190
POSI	TEMPLATE	90	17.2114	6.76181	.71276
	TRIAL-and-ERROR	17	32.1324	9.65139	2.34081
	Total	107	19.5820	9.08198	.87799

		Sum of Squares	df	Mean Square	F	Sig.
REP	Between Groups	1547.396	1	1547.396	12.702	.001
	Within Groups	12791.553	105	121.824		
	Total	14338.949	106			
AVBLE	Between Groups	1574.049	1	1574.049	15.623	.000
	Within Groups	10578.968	105	100.752		
	Total	12153.017	106			
AA	Between Groups	8013.093	1	8013.093	100.507	.000
	Within Groups	8371.272	105	79.726		
	Total	16384.365	106			
POSI	Between Groups	3183.480	1	3183.480	60.123	.000
	Within Groups	5559.651	105	52.949		
	Total	8743.131	106			

Table 5-24 Cluster analysis of customer relationship problem sense making information search heuristic orientation

Descriptive Statistics

	Mean	Std. Deviation	N
Customer Relationship Problem Sense-Making Enhancements	6.3730	1.19209	107
Quality on DW Customer Relationship Data for Problem Enacment	4.5409	.79681	107

Correlations

		Customer Relationship Problem Sense-Making Enhancements	Quality on DW Customer Relationship Data for Problem Enacment
Customer Relationship Problem Sense-Making Enhancements	Pearson Correlation	1	.530**
	Sig. (2-tailed)	.	.000
	N	107	107
Quality on DW Customer Relationship Data for Problem Enacment	Pearson Correlation	.530**	1
	Sig. (2-tailed)	.000	.
	N	107	107

**. Correlation is significant at the 0.01 level (2-tailed).

Table 5-25 Correlation between QUALITY and ENACENH

The next step was to test the direction and strength of the association. Multivariate regression analysis was conducted for that purpose in both directions. Table 5-26 shows the multivariate regression analysis with the cognitive construct ENACENH as the dependent variable. The unstandarized coefficient beta indicates that when QUALITY goes up by 1, then ENACENH goes up by .793 (p <.000). Having just one independent vaiable the standardized beta is the correlation. Plotting the studentized residuals against the standardized predicted values did not show a consistent pattern. This means equal variances (i.e. no issues with heteroscedasticity) [Hair Jr. *et al.* 1998]. Therefore, I concluded that ENACENH is a dependent variable of QUALITY.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.530[a]	.281	.274	1.01570

a. Predictors: (Constant), Quality on DW Customer Relationship Data for Problem Enacment

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	42.312	1	42.312	41.014	.000[a]
	Residual	108.323	105	1.032		
	Total	150.634	106			

a. Predictors: (Constant), Quality on DW Customer Relationship Data for Problem Enacment

b. Dependent Variable: Customer Relationship Problem Sense-Making Enhancements

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2.772	.571		4.858	.000		
	Quality on DW Customer Relationship Data for Problem Enacment	.793	.124	.530	6.404	.000	1.000	1.000

a. Dependent Variable: Customer Relationship Problem Sense-Making Enhancements

Table 5-26 ENACENH as the dependent variable of QUALITY

Table 5-27 shows the multivariate regression analysis with QUALITY as the dependent variable. Plotting the studentized residuals against the standardized predicted values did not show a consistent pattern. This means equal variances (i.e. no issues with heteroscedasticity) [Hair Jr. *et al.* 1998]. The unstandarized coefficient beta indicates that when ENACENH goes up by 1, then QUALITY goes up by .354 (p <.000). However, the low unstandarized coefficient beta does not provide conclusive evidence that ENACENH is a dependent variable of QUALITY.

Overall this means that there is an association between quality on data warehouse customer relationship data for problem enactment (QUALITY) and customer relationship problem sense making enhancements (ENACENH). This suggests strong support for hypothesis H4.1. Furthermore, there is a positive association and there is strong evidence supporting ENACENH and the dependent variable and QUALITY and the independent variable in this association.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.530[a]	.281	.274	.67891

a. Predictors: (Constant), Customer Relationship Problem Sense-Making Enhancements

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	18.904	1	18.904	41.014	.000[a]
	Residual	48.397	105	.461		
	Total	67.301	106			

a. Predictors: (Constant), Customer Relationship Problem Sense-Making Enhancements

b. Dependent Variable: Quality on DW Customer Relationship Data for Problem Enacment

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2.283	.359		6.367	.000		
	Customer Relationship Problem Sense-Making Enhancements	.354	.055	.530	6.404	.000	1.000	1.000

a. Dependent Variable: Quality on DW Customer Relationship Data for Problem Enacment

Table 5-27 QUALITY as the dependent variable of ENACENH

In order to have further confidence on the support to the association between QUALITY and ENACENH, I conducted a confirmatory analysis using structural equation modeling. As a result of the research findings in phase I, the information search behavior constructs about the modes and heuristics were clustered in two cluster solutions (i.e. FOCUSMO vs. SCANMO and TEMPLATE vs. TRIAL-and-ERROR). Therefore, it was not feasible to include these clusters as part of a confirmatory analysis using structural equation modeling because these variables are not metric [Hair Jr. *et al.* 1998]. Also, the strong support to H4.1 with ENACENH as the dependent variable and QUALITY as the independent variable suggests a recursive model with QUALITY as the exogenous variable and ENACENH as the endogenous variable.

From a confirmatory perspective I will focus on the recursive part of the research model that covers the association between QUALITY and ENACENH. This model is illustrated in APPENDIX J: DETAILED MODEL TO BE TESTED USING STRUCTURAL EQUATION MODELING. I entered this model in AMOS and executed it against the sample, N=107.

AMOS reported that the detailed model was overinformed by 525 degrees of freedom (see Figure 5.4). The good news is that this means that the detailed model could be tested using AMOS. Unfortunately, the bad news is that using the guidelines of five cases per parameter [Hair Jr. *et al.* 1998] requires 345 cases and my sample only has N=107. Furthermore, the bootstrap procedure is not applicable when the sample size is too small for the requirements of the model ([Yung and Bentler 1996] p. 223).

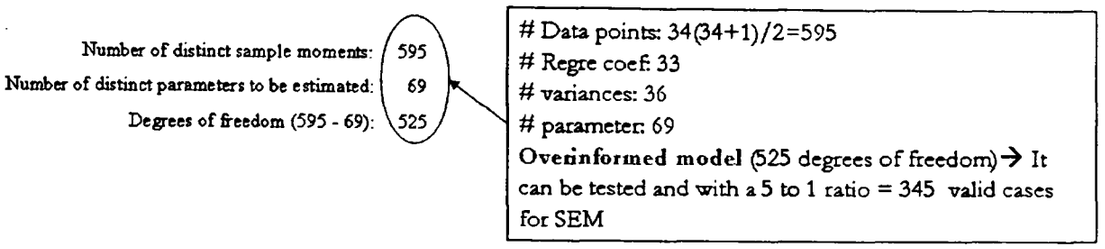


Figure 5.4 Degrees of freedom for the detailed model

So, I decided to proceed with the following reductionist approach:

- I considered all the constructs of first order in the detailed model specification as observed variables. In this way I reduced the amount of parameters.
- I applied confirmatory factor analysis using the multivariate structural equation modeling technique following the guidelines in [Anderson and Gerbing 1988, Hair Jr. *et al.* 1998, Harris and Schaubrock 1990].

I entered this reduced model in AMOS and executed it against the sample, N=107. AMOS reported that the detailed model was overinformed by 12 degrees of freedom (see Figure 5.5). This means that, as expected, the reduced model can be tested, and will require 80 cases (my sample has N=107) using the guidelines of five cases per parameter [Hair Jr. *et al.* 1998].

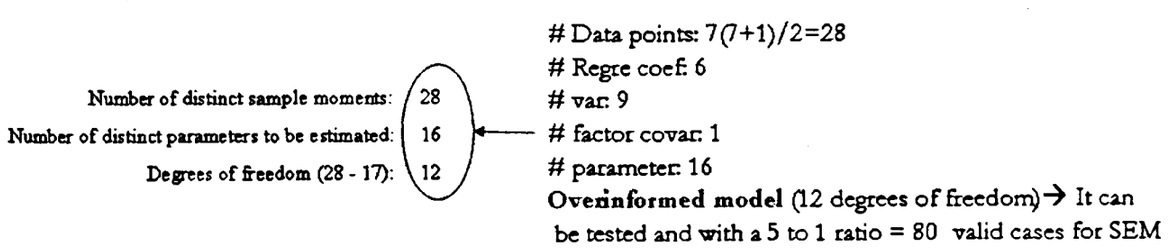


Figure 5.5 Degrees of freedom for the reduced model

This reduced model and the details about the results are in APPENDIX K: REDUCED MODEL TESTED USING STRUCTURAL EQUATION MODELING.

In summary, AMOS reported a beta standardized regression coefficient .871 (R^2 = 75.9%) indicating that the strength of the association is high [Hair Jr. *et al.* 2003].

Models tested under SEM cannot be accepted, they just cannot be rejected as recommended in [Anderson and Gerbing 1988]. The goodness-of-fit analysis did not reject the null hypothesis Ho: "The sample observations and the predicted estimates by the model have differences." On the contrary, all the fit indexes indicated model fit. The model misspecification analysis suggested that modifications would not be needed. I concluded that any further incorporation of parameters into the model would result in an overfitted model. Indeed, when an initial model fits well, it is probably unwise to modify it to achieve even better fit because modifications may simply be fitting small idiosyncratic characteristics of the sample [Byrne 2001].

The reliability of the latent constructs was computed using the construct reliability and the variance extracted measures [Hair Jr. *et al.* 1998]. ENACENH exhibits acceptable construct reliability (.82) and variance extracted (.70). Unfortunately, QUALITY has marginal construct reliability (.44) and unacceptable variance extracted (.32). Therefore, I cannot consider the results of the confirmatory analysis (i.e. the strength of the association is high) as conclusive.

5.8 PHASE III: CONFIRMATORY ANALYSIS OF ASSOCIATION BETWEEN COGNITIVE/ENVIRONMENTAL AND BEHAVIORAL CONSTRUCTS

5.8.1 Confirmatory Analysis of Association Between Cognitive/Environmental Constructs and Behavioral Search Patterns

The expectation is that there is an association between customer relationship problem sense making information search behavioral patterns MODE with customer relationship problem sense making enhancements ENACENH and with the quality of data warehouse customer relationship data for problem enactment QUALITY. In phase II, I found that FOCUSMO and SCANMO are two different behavioral patterns in MODE (i.e. hypothesis H3.1). Discriminant analysis was conducted for that purpose. A requirement for discriminant analyses is that clusters should have equal size [Hair Jr. *et al.* 2003]. In this research the sizes are FOCUSMO (75 cases) and

SCANMO (32 cases). Therefore, in order to satisfy the equality in size requirement, I randomly selected 32 cases in FOCUSMO.

Discriminant power of cognitive and quality constructs over customer relationship problem sense making information search mode orientations (see Table 5-28) shows the discriminant analysis with clusters FOCUSMO and SCANMO as the dependent variable. The analysis identified the cognitive first order SELFE variable and the quality first order RICH variable as discriminating FOCUSMO vs. SCANMO with a statistical significant (p = .001) predicting power of 65.6% of hit rate which falls slightly below the recommended 70% [Hair Jr. *et al.* 2003].

Structure Matrix

	Function 1
Customer Relationship Problem Enacting Self-Efficacy	.796
Customer Relationship Problem Enactment Knowledge Enhacement [a]	.302
Data Richness	-.229
Competitor Insights [a]	.146
Data Equivocality [a]	-.103
Customer Insights [a]	.076
Data Integration [a]	.034

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function

a. This variable not used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.801	13.543	2	.001

Classification Results[a]

			Predicted Group Membership		
		Ward Method	1	2	Total
Original	Count	Focus Orientation	20	12	32
		Scan Orientation	10	22	32
	%	Focus Orientation	62.5	37.5	100.0
		Scan Orientation	31.3	68.8	100.0

a. 65.6% of original grouped cases correctly classified.

Table 5-28 SELFE and RICH as discriminant predictors of mode orientations

Further analysis of the means of SELFE and RICH (see Table 5-29) indicates that there is not much difference in means of RICH between the two clusters leaving all the discriminating contribution to SELFE. I decided not to consider RICH as discriminant of mode orientations. Therefore, I concluded that SELFE is a discriminant of customer relationship problem sense making information search mode orientations. Given that the coefficient of SELFE is positive in the discriminant function, then the higher value in the SCANMO orientation (see the .491 value in group centroids in Table 5-29) indicates that high scores in SELFE will lead to a higher likelihood of predicting the SCANMO mode orientation and the low scores in SELFE will lead to a higher likelihood of predicting the FOCUSMO mode orientation. This means that there is an association between customer relationship

problem sense making information search behavioral mode patterns MODE and customer relationship problem sense making enhancements ENACENH (SELFE is a dimension in ENACENH). This suggests strong support for hypothesis H4.2. However, discriminant analysis did not find any dependence between information search behavioral mode patterns MODE and the quality of data warehouse customer relationship data for problem enactment QUALITY. This suggests that hypothesis H4.4 is not supported.

Cluster		Mean	Std. Deviation	Valid N (listwise)
Focus Orientation	Data Richness	5.9977	.88128	32
	Customer Relationship Problem Enacting Self-Efficacy	6.5480	1.41402	32
Scan Orientation	Data Richness	5.7714	1.11906	32
	Customer Relationship Problem Enacting Self-Efficacy	7.6172	1.32217	32
Total	Data Richness	5.8845	1.00567	64
	Customer Relationship Problem Enacting Self-Efficacy	7.0826	1.46096	64

Functions at Group Centroids

	Function
Ward Method	1
Focus orientation	-.491
Scan Orientation	.491

Unstandardized canonical discriminant functions evaluated at group means

Table 5-29 Group centroids discriminating FOCUSMO vs. SCANMO

5.8.2 Confirmatory Analysis of Association between Cognitive/Environmental Constructs and Behavioral Heuristic Patterns

The expectation is that there is an association between customer relationship problem sense making information heuristic behavioral patterns HEUR with customer relationship problem sense making enhancements ENACENH and with the quality of data warehouse customer relationship data for problem enactment QUALITY. In phase II, I found that TEMPLATE and TRIAL-and-ERROR are two different behavioral patterns in HEUR (i.e. hypothesis H3.2). Discriminant analysis was conducted for that purpose. A requirement for discriminant analyses is that clusters should have equal size [Hair Jr. *et al.* 2003]. In this research the sizes are TEMPLATE (93 cases) and TRIAL-and-ERROR (14 cases). Therefore, in order to satisfy the equality in size requirement, I randomly selected 14 cases in TEMPLATE.

Discriminant power of cognitive and quality constructs over customer relationship problem sense making information search heuristic orientations (see Table 5-30) shows the discriminant analysis with clusters TEMPLATE and TRIAL-and-ERROR as the dependent variable. The analysis identified the cognitive second order TASKUT variable as discriminating TEMPLATE vs. TRIAL-and-ERROR with a statistical significant ($p < .000$) predicting power of 85.7% of hit rate which exceeds the recommended 70% [Hair Jr. *et al.* 2003].

Structure Matrix

	Function
	1
Data Task Utility	1.000
Customer Relationship Problem Sense-Making Enhancements	.472
Data Enactment Utility	.376

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function

a. This variable not used in the analysis.

Classification Results[a]

			Predicted Group Membership		
		Ward Method	1	2	Total
Original	Count	Template Orientation	12	2	14
		Trial-and-error Orientation	2	12	14
	%	Template Orientation	85.7	14.3	100.0
		Trial-and-error Orientation	14.3	85.7	100.0

a. 85.7% of original grouped cases correctly classified.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.433	21.342	1	.000

Functions at Group Centroids

	Function
Ward Method	1
Template Orientation	1.103
Trial-and-error orientation	-1.103

Unstandardized canonical discriminant functions evaluated at group means

Table 5-30 TASKUT as discriminant predictor of heuristic orientations

Therefore, I concluded that TASKUT is a discriminant of customer relationship problem sense making information search heuristic orientations. Given that the coefficient of TASKUT is positive in the discriminant function, then the higher value in the TEMPLATE orientation (see the 1.103 value in group centroids in Table 5-30) indicates that high scores in TASKUT will lead to a higher likelihood of predicting the TEMPLATE heuristic orientation and the low scores in TASKUT will lead to a higher likelihood of predicting the TRIAL-and-ERROR heuristic orientation.

In phase I, the intensity of the data warehouse anchoring and adjustment search heuristic (AA) had departures from normality. This has no effect on cluster analysis but it is controversial in discriminant analysis [Hair Jr. *et al.* 1998]. However: (i) my discriminant analysis did not consider AA as the dependent variable but a cluster with

the four heuristics included AA. Either way, AA for the randomly selected cases for the discriminant analysis (i.e. in order to have same cluster size) did not have normality issues in AA (Skewness= 0.929, Kurtosis=0.216).

This means that there is an association between customer relationship problem sense making information search behavioral heuristic patterns HEUR and the quality of data warehouse customer relationship data for problem enactment QUALITY (TASKUT is a dimension in QUALITY). This suggests strong support for hypothesis H4.5. However, discriminant analysis did not find any dependence between information search behavioral heuristic patterns HEUR and customer relationship problem sense making enhancements ENACENH. This suggests that hypothesis H4.3 is not supported.

5.9 PHASE IV: EXPLORATORY ANALYSIS OF MODERATORS IMPACT ON THE ASSOCIATION BETWEEN ENVIRONMENTAL AND COGNITIVE CONSTRUCTS

5.9.1 Analysis of Moderating Impact of CRM Job Function

In general terms, a moderator is a qualitative (e.g. sex, race, class) or quantitative (e.g. level of reward) variable that affects the direction and/or strength of the relation between an independent or predictor variable and a dependent or criterion variable. Specifically within a correlational analysis framework, a moderator is a third variable that affects the zero-order correlation between two other variables. In the more familiar analysis of variance (ANOVA) terms, a basic moderator effect can be represented as an interaction between a focal independent variable and a factor that specifies the appropriate conditions for its operation [Baron and Kenny 1986]. There were no clear a-priori expectations with respect to CRM job function (JOBFUNC) as a moderator on the association between QUALITY and ENACENH.

The variable CRM job function (JOBFUNC) introduces four groups in QUALITY and ENACENH (sales support, marketing support, customer service support, and other CRM functions). The first step in the analysis of the moderating impact consists of testing the homogeneity of variance for the dependent variable across all level combinations. The null hypothesis is that the error variance in the dependent

Descriptive Statistics

Dependent Variable: Customer Relationship Problem Sense-Making Enhancements

61 What is the primary	Mean	Std. Deviation	N
Sales Support	5.9173	1.33560	21
Marketing Support	6.5768	1.07644	24
Customer Service Support	6.5373	1.10542	21
Other CRM Functions	6.6494	1.25341	16
Total	6.4119	1.20367	82

Levene's Test of Equality of Error Variances

Dependent Variable: Customer Relationship Problem Sense-Making Enhancements

F	df1	df2	Sig.
.788	3	78	.504

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+item_61 * quality+item_61+quality

Descriptive Statistics

Dependent Variable: The Quality of DW Customer Relationship Data for Problem Enactment

61 What is the primary	Mean	Std. Deviation	N
Sales Support	4.4059	.92704	21
Marketing Support	4.4861	.80948	24
Customer Service Support	4.6380	.86883	21
Other CRM Functions	4.7708	.64451	16
Total	4.5600	.82444	82

Levene's Test of Equality of Error Variances(a)

Dependent Variable: The Quality of DW Customer Relationship Data for Problem Enactment

F	df1	df2	Sig.
1.119	3	78	.346

Tests the null hypothesis that the error variance o the dependent variable is equal across groups.

a. Design: Intercept+item_61 * enacenh+item_61+enacenh

Table 5-31 Homogeneity of variance for QUALITY and ENACENH across all level combinations produced by JOBFUNC

Tests of Between-Subjects Effects

Dependent Variable: Customer Relationship Problem Sense-Making Enhancements

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	40.730[b]	7	5.819	5.619	.000	.347
Intercept	13.249	1	13.249	12.795	.001	.147
item_61 * quality	2.252	3	.751	.725	.540	.029
item_61	2.257	3	.752	.726	.539	.029
quality	32.490	1	32.490	31.377	.000	.298
Error	76.624	74	1.035			
Total	3488.624	82				
Corrected Total	117.354	81				

a. Computed using alpha = .05

b. R Squared = .347 (Adjusted R Squared = .285)

Dependent Variable: The Quality of DW Customer Relationship Data for Problem Enactment

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	16.879[b]	7	2.411	4.674	.000	.307
Intercept	11.846	1	11.846	22.962	.000	.237
item_61 * enacenh	.133	3	.044	.086	.968	.003
item_61	.247	3	.082	.160	.923	.006
enacenh	14.904	1	14.904	28.890	.000	.281
Error	38.177	74	.516			
Total	1760.160	82				
Corrected Total	55.055	81				

a. Computed using alpha = .05

b. R Squared = .307 (Adjusted R Squared = .241)

Table 5-32 ANOVA analyses for the iteration of JOBFUNC and the independent variable on the dependent variable

variable is equal across the groups. Lack of statistical significance implies that you cannot reject the null hypothesis (i.e. equality of variances through the groups). This analysis can be done with Levene's test and has to be applied to both QUALITY and ENACENH. Levene's test (see Table 5-31) exhibits lack of significance for ENACENH (p = .504) and for QUALITY (p = .346). So, there is no reason to believe that the equal variances assumption is violated. Thus, the differences in the group standard deviations observed in the descriptive statistics tables in Table 5-31 are due to random variation.

Once the homogeneity of variance had been confirmed then ANOVA analysis was conducted as suggested in [Baron and Kenny 1986] for both QUALITY and ENACENH. Lack of statistical significance for the interaction term <dependent_variable * moderator_variable> implies that you cannot reject the null hypothesis (i.e. no interaction, differences in group means are due to random variation).

JOBFUNC was implemented in the questionnaire as item_61. ANOVA analyses (see Table 5-32) exhibited lack of statistical significance for the interaction term QUALITY * item_61 (p = .540) and the interaction term ENACENH * item_61 (p = .968). This implies that you cannot reject the null hypothesis (i.e. no interaction, differences in group means are due to random variation). So, there is no reason to believe that the lack of interaction assumption is violated. Thus, again, the differences in group means observed in the descriptive statistics tables in Table 5-31 are due to random variation [Chin, Marcolin and Newsted 1996]. Therefore, these research results do not support hypothesis H5.1

5.9.2 Analysis of Moderating Impact of the Number of Supported CRM Data Warehouse Functions

The moderator variable number of supported CRM data warehouse functions (DWFUNC) has a multiple answer format. All the combinations $\sum_{n=1}^{4}\binom{4}{n}=15$ where $\binom{4}{n}=\frac{4!}{n!(4-n)!}$ were present in the sample. I considered that the most interesting

groups were the group that I labeled as "just one CRM function supported by the data warehouse" and the group that I labeled as "more than one CRM function supported by the data warehouse". I implemented a dummy variable [Hair Jr. *et al.* 2003], labeled dummy60, in order to create these two groups in QUALITY and ENACENH. There was a clear expectation of support for hypothesis H6.1. Like with the JOBFUNC variable, Levene's test was applied to both QUALITY and ENACENH.

Levene's test (see Table 5-33) exhibits a lack of significance for ENACENH (p = .422) and for QUALITY (p = .118). So, there is no reason to believe that the equal variances assumption is violated. Thus, the differences in group standard deviations observed in the descriptive statistics tables in Table 5-33 are due to random variation. Once the homogeneity of variance had been confirmed then ANOVA analysis was conducted as suggested in [Baron and Kenny 1986] for both QUALITY and ENACENH.

ANOVA analyses (see Table 5-34) exhibited lack of statistical significance for the interaction term QUALITY * dummy60 (p = .410) and the interaction term ENACENH * item_61 (p = .251). This implies that you cannot reject the null hypothesis (i.e. no interaction, differences in group means are due to random variation). So, there is no reason to believe that the lack of interaction assumption is violated. Thus, again, the differences in the group means observed in the descriptive statistics tables in Table 5-33 are due to random variation [Chin *et al.* 1996]. Therefore, these research results do not support hypothesis H6.1

5.10 PHASE V: EXPLORATORY ANALYSIS OF GROUP DIFFERENCES

The impact analysis of JOBFUNC and DWFUNC as producers of group differences should be explored over the whole model represented in Figure 3.1. Notice that the language in hypotheses H5.2, H5.3, H5.4, H5.5, H6.2, H6.3, H6.4, and H6.5 focuses on group differences while the language in H5.1and H6.1 focuses on the association between variables.

Descriptive Statistics

Dependent Variable: Customer Relationship Problem Sense-Making Enhancements

60 DUMMY Which CRM	Mean	Std. Deviation	N
Just one CRM function	6.5132	1.34599	12
More than one CRM Function	6.5000	.93385	12
Total	6.5066	1.13296	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Customer Relationship Problem Sense-Making Enhancements

F	df1	df2	Sig.
.669	1	22	.422

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+dummy60 * quality+dummy60+quality

Descriptive Statistics

Dependent Variable: The Quality of DW Customer Relationship Data for Problem Enactment

60 DUMMY Which CRM	Mean	Std. Deviation	N
Just one CRM function	4.5944	1.03297	12
More than one CRM Function	4.6257	.57362	12
Total	4.6101	.81728	24

Levene's Test of Equality of Error Variances(a)

Dependent Variable: The Quality of DW Customer Relationship Data for Problem Enactment

F	df1	df2	Sig.
2.651	1	22	.118

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+dummy60 * enacenh+dummy60+enacenh

Table 5-33 Homogeneity of variance for QUALITY and ENACENH across all level combinations produced by dummy60

Tests of Between-Subjects Effects

Dependent Variable: Customer Relationship Problem Sense-Making Enhancements

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	6.016[b]	3	2.005	1.706	.198	.204
Intercept	10.290	1	10.290	8.755	.008	.304
dummy60 * quality	.833	1	.833	.708	.410	.034
dummy60	.795	1	.795	.676	.421	.033
quality	2.103	1	2.103	1.789	.196	.082
Error	23.507	20	1.175			
Total	1045.576	24				
Corrected Total	29.523	23				

a. Computed using alpha = .05

b. R Squared = .204 (Adjusted R Squared = .084)

Dependent Variable: The Quality of DW Customer Relationship Data for Problem Enactment

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	3.528[b]	3	1.176	1.988	.148	.230
Intercept	5.553	1	5.553	9.385	.006	.319
dummy60 * enacenh	.827	1	.827	1.397	.251	.065
dummy60	.831	1	.831	1.404	.250	.066
enacenh	1.487	1	1.487	2.514	.129	.112
Error	11.834	20	.592			
Total	525.431	24				
Corrected Total	15.363	23				

a. Computed using alpha = .05

b. R Squared = .230 (Adjusted R Squared = .114)

Table 5-34 ANOVA analyses for the iteration of DWFUNC and the independent variable on the dependent variable

This means that the impact of JOBFUNC and DWFUNC as producers of group differences will be explored as a factor that introduces group differences in terms of means. The a-priori expectations were that Hypotheses H5.3, H5.4, H5.5, H6.2, H6.3, H6.4, and H6.5 would be supported. However, these expectations were not so clear after the research findings in phase IV, where hypotheses H5.1 and H6.1 were tested, indicated a lack of support. Fourteen ANOVA analyses were conducted.

In relation to JOBFUNC, Levene's test for the seven variables QUALITY, ENACENH, SCAN/FOCUS, REP, AVBLE, AA and POSI was not significant.

So, there is no reason to believe that the equal variances assumption was violated. Thus, the differences in group standard deviations observed in the descriptive statistics tables were due to random variation. ANOVA analyses exhibited lack of statistical significance for the term item_61 in the seven variables. This implies that you cannot reject the null hypothesis (i.e. no difference in means across the groups). So, there is no reason to believe that the lack of differences assumption is violated. Thus, the differences in group means observed in the descriptive statistics were due to random variation. Therefore, these research results do not support hypothesis H5.2, H5.3, nor H5.4. Hypothesis H5.5 involves four heuristics (REP, AVBLE, AA, POSI). AA had departures from normality, therefore, H5.5 should be cautiously interpreted with respect to this heuristic. Overall, H5.5 is not supported with the comment that the anchor and adjustment heuristic had departures for normality.

I implemented a dummy variable [Hair Jr. *et al.* 2003], labeled dummy60, in order to explore the most interesting groups in the number of supported CRM data warehouse functions (DWFUNC). This dummy variable had two groups (i.e. "just one CRM function supported by the data warehouse" and "more than one CRM function supported by the data warehouse"). Levene's test for the six out of the seven variables QUALITY, ENACENH, SCAN/FOCUS, REP, AVBLE, AA and POSI was not significant. The exception was QUALITY (i.e. Levene was significant). This significance in the variances across the two groups came from the data task utility dimension TASKUT (i.e. Levene was not significant in the data enactment dimension ENACUT). Because the group's means are not significantly different in QUALITY

(4.5944 vs. 4.6257), I considered that this violation of the homogeinity of variance was not an issue.

ANOVA analyses exhibited lack of statistical significance for the term dymmy60 in the seven variables. This implies that you cannot reject the null hypothesis (i.e. no difference in means across the groups). So, there is no reason to believe that the lack of differences assumption is violated. Thus, the differences in group means observed in the descriptive statistics were due to random variation. Therefore, these research results do not support hypothesis H6.2, H6.3, nor H6.4. Hypothesis H6.5 involves four heuristics (REP, AVBLE, AA, POSI). AA had departures from normality, therefore, H6.5 should be cautiously interpreted with respect this heuristic. Overall, H6.5 is not supported with the comment that the anchor and adjustment heuristic had departures for normality.

Summarizing, ANOVA analyses (see Table 5-35) exhibited a lack of statistical significance for the F statistic in all the variables for both JOBFUNC and DWFUNC. This implies that you cannot reject the null hypothesis (i.e. differences in group means are due to random variation). So, there is no reason to believe that the lack of differences in group means assumption is violated. Thus, again, the differences in group means observed in the descriptive statistics tables are due to random variation [Chin *et al.* 1996].

TYPE OF CONSTRUCT	CONSTRUCT	JOBFUNC	DWFUNC
		F *p*	F *p*
Environmental	QUALITY	0.547	0.928
Cognitive	ENACENH	0.184	0.978
Behavioral	SCAN/FOCUS	0.716	0.195
	REP	0.795	0.205
	AVBLE	0.222	0.268
	AA	0.408	0.14
	POSI	0.711	0.539

Table 5-35 ANOVA significance levels for the moderating impact of JOBFUNC and DWFUNC in variables involved in associations

5.11 PHASE VI: CONFIRMATORY AND EXPLANATORY ANALYSIS OF RESEARCH FINDINGS

I conducted a qualitative analysis pursuing confirmatory and explanatory objectives with the former seeking validation of the quantitative research findings, and the latter seeking relevant implications for organizations (e.g. [Myers 1997]). Theoretical implications of quantitative research findings were not in the scope of this qualitative analysis. Also, research findings #7 and #8 were excluded from this analysis due to my concerns with respect to the lack of expertise of interviewees in discriminant concepts. The data collection technique was semi-structured interviews (e.g. [Remenyi *et al.* 1998, Hair Jr. *et al.* 2003]) and the qualitative analysis consisted of the following steps: data reduction, display, and conclusion (e.g. [Marshall and Rossman 1995, Miles and Huberman 1994]). Data reduction followed a protocol which consisted of the coding and classification of answers and a filing system for the interviews. The code system was the following: <answer to open question> *is allocated to* <Practical Implication theme code > which is allocated to <Quantitative research finding code>. In summary, each research finding required one structured question (seeking validation) and a series of pre-determined open questions (seeking practical implications). See APPENDIX M: RESEARCH FINDINGS VALIDATION SEMI-STRUCTURED INTERVIEWS for details.

A pattern emerged after the third interview and it was confirmed in a fourth interview. I assessed that with these four interviews there was no need for further interviews. Participants were selected from my personal network of professional contacts. The first interviewee is based in San Diego, USA, and he has extensive experience as a data warehouse and CRM consultant in the retail industry. The second interviewee is based in Buenos Aires, Argentina, and he works in a major retail bank as a data warehouse practitioner supporting requests from CRM end users. The third interviewee is based in Zurich, Switzerland, and he has extensive experience as a data warehouse and CRM practitioner in several industries (now in a mobile communications operator). The fourth interviewee is based in London, UK, and he has extensive experience as a CRM consultant in several industries and holds a PhD degree (his research was on perceived information value and information quality).

Responses to the structured questions were summarized with their means. Overall, there was an overwhelmingly strong support to all the research findings with the only exception in finding #9 with respect to the number of supported CRM data warehouse functions (see Table 6 -6).

Although the participants did not object to the finding that DWFUNC is not a moderator, their support of this finding was small indicating a sharp confrontation with expectations. Therefore, I decided to apply the analysis that I did in phase V to most of QUALITY constructs that were not explored, that is, TASKUT, ENACUT, CUST, COMP, INT, RICH, and EQUI (see APPENDIX N: ADDITIONAL ANOVA ANALYSES TESTING GROUPS DIFFERENCES ON QUALITY VARIABLES FACTORED BY DWFUNC).

ANOVA analyses (see Table 5-36) exhibited a lack of statistical significance for the F statistic in all the variables for DWFUNC. The reported partial eta squared statistic for such terms were insignificant. This implies that you cannot reject the null hypothesis (i.e. differences in group means are due to random variation). So, there is no reason to believe that the lack of differences in the group means assumption is violated. Thus, again, the differences in the group means observed in the descriptive statistics tables are due to random variation [Chin *et al.* 1996]. Therefore, DWFUNC is not a producer of group differences with respect to the indicators of QUALITY.

5.12 GENERALIZATION OF RESEARCH FINDINGS

5.12.1 Key Aspects for a Critical Assessment

External generalization is necessary in order to claim that the research findings can be generalized. This critical assessment is presented in the Generalization of Research Findings Assessment section. However, before making such an assessment the following aspects should be considered when assessing whether or not research findings can be considered as candidates for external generalization to the population (e.g. [Schwab 1999]): (i) The sample frame, (ii) the reliability of the sample and validity of the research findings and (iii) the size of the sample size. The sample

frame assessment should include the representativeness of the sample and the normality.

The following includes a section addressing each of the three mentioned aspects and a short final section with the overall assessment about the generalization of the research findings. All the data in the following sections come from the analyses that have already been made (i.e. there is no new data), however the presentation frame is different as it focuses on the assessment of generalization.

5.12.2 Sample Frame Assessment

The representativeness of the sample was successfully assessed in the REPRESENTATIVENESS OF THE SAMPLE DATA section concluding that the sample was representative of the population. Figure 5.6 contains the statistics assessing normality for all the scales involved in the research findings. Overall, scales exhibited normality with just the exception of the intensity of the data warehousing anchoring and adjustment search heuristic (AA) that had departures from normality. This exception had no impact on most of the research findings because the techniques used, that is cluster analysis and, to a lesser extent, discriminant analysis, do not have normality as a requirement. Nevertheless, the clusters considered exhibited normality for AA. However, ANOVA results for hypotheses H5.5 and H6.5 are reported commenting this issue. Overall, normality was not an issue for the scales involved in the research findings.

5.12.3 Reliability and Validity Assessment

The reliability of the representativeness of the sample was successfully assessed using several statistics (see Figure 5.6). Overall, reliability was not an issue for the scales involved in the research findings. Construct validity was assessed empirically for the environmental and cognitive constructs testing the unidimensionality of the involved constructs. Information search behavior constructs have summated scales which are not suitable for unidimensionality analysis.

Unidimensionality is exhibited in CUST, COMP, INT, RICH, and EQUI with respect to QUALITY in factor analysis. Factor loadings for the items rank between .71 and

.90 which is considered high and a few fall in the rank .41 to .70 which is considered moderate [Hair Jr. *et al.* 2003]. The variance extracted R^2 =69.646 exceeds the cut-off of .5 (see Table 5-13).

Unidimensionality is exhibited in TASKUT and ENACUT with respect to QUALITY in factor analysis:

- Factor analysis of summates scales: Loadings rank between .71 and .90 which is considered high. EQUI should also be considered high. INT falls in the rank .41 to .70 which is considered moderate [Hair Jr. *et al.* 2003]. The variance extracted R^2 =67.741 exceeds the cut-off of .5 (see Table 5-15).
- Factor analysis of items: The variance extracted R^2 =47.800 falls slightly below .5 (see Table 5-14).

Unidimensionality is exhibited in KNOWENH and SEFE with respect to ENACENH in factor analysis:

- Factor analysis of items: Loadings rank between .71 and .90 which is considered high and few fall in the rank .41 to .70 which is considered moderate [Hair Jr. *et al.* 2003]. The variance extracted R^2 =58.376 exceeds the cut-off of .5 (see Table 5-20).

Discriminant validity is exhibited in CUST, COMP, INT, RICH and EQUI (see Table 5-13) and in KNOWENH, and SELFE (see Table 5-20) in factor analysis. Most of the cross-loadings on other latent constructs are smaller than .30, which is considered small, and only three fall in the rank .41 to .70, which is considered moderate [Hair Jr. *et al.* 2003]. Overall, divergent validity is convincingly exhibited in QUALITY, CUST, COMP, INT, RICH, EQUI, KNOWENH, SELFE, TASKUT and ENACUT.

Convergent validity x-X, also known as criterion validity [Schutt 1996], concurrent validity [Reaves 1992], and predictive validity [Rafilson 1999], is moderately exhibited in (i) QUALITY and its dimensions TASKUT and ENACUT and in (ii) ENACENH and its dimensions KNOWENH and SELFE. Convergent validity is the confirmation of the validity of the scale x as a measure of the construct X by comparing the scores of x with the results yielded at the same time (i.e. concurrently) by another scale y [Easterby-Smith *et al.* 1991] and finding a high correlation

TYPE OF CONSTRUCT	CONSTRUCT	DWFUNC F p
Environmental	CUST	0.693
	COMP	0.545
	TASKUT	0.834
	INT	0.512
	RICH	0.517
	EQUI	0.309
	ENACUT	0.604
	QUALITY	0.928

Table 5-36 ANOVA significance levels for the moderating impact of DWFUNC in quality variables

	SCALE	Alpha >.70 (>.60 for exploratory research) item-to-total correlations > .5 and inter-items correlation > .3		S in +/- 1 and K in +/- 3			
		Before purification/outliers		After Purification and Outliers Removal			
		Alpha	# Items	Alpha	# Items	Z Skewness	Z Kurtosis
SCAN/FOCUS	Intensity of data warehouse scanning	0.66	2			0.2253	-0.7993
REP	Intensity of data warehouse representativeness search heuristic	0.66	4			0.4615	0.0887
AVBLE	Intensity of data warehouse availability search heuristic	0.67	4			0.7102	0.7099
AA	Intensity of data warehouse anchoring and adjustment search heuristic	0.78	4			1.4583	3.3169
POSI	Intensity of data warehouse positivy search heuristic	0.64	4			0.9589	1.0234

	SCALE	Alpha >.70 (>.60 for exploratory research) item-to-total correlations > .5 and inter-items correlation > .3			Variance extracted > .5	S in +/- 1 and K in +/- 3	
		After Factor Analysis					
		Alpha	Composite Construct Reliability	# Items	Variance extracted	Z Skewness	Z Kurtosis
KNOWENH	Customer Relationship Problem Enactment Knowledge Enhancement	0.90	NA	7	NA	-0.7846	0.2283
LSELFE	Low Customer Relationship Problem Enacting Self-Efficacy	0.82	NA	4	NA	-0.7222	0.1426
MSELFE	Medium Customer Relationship Problem Enacting Self-Efficacy	0.84	NA	4	NA	-0.5601	0.3043
HSELFE	High Customer Relationship Problem Enacting Self-Efficacy	0.80	NA	4	NA	-0.3638	-0.0091
SELFE	Customer Relationship Problem Enacting Self-Efficacy	0.92	NA	12	NA	-0.3695	-0.3634
ENACENH	Customer Relationship Problem Sense-Making Enhancements	0.94	NA	19	0.58376	-0.4294	-0.3026

	SCALE	Alpha >.70 (>.60 for exploratory research) item-to-total correlations > .5 and inter-items correlation > .3			Variance extracted > .5	S in +/- 1 and K in +/- 3	
		After Factor Analysis					
		Alpha	Composite Construct Reliability	# Items	Variance extracted	Z Skewness	Z Kurtosis
CUST	Customer Insights	0.82	NA	2	NA	0.0390	-1.3241
COMP	Competitor Insights	0.71	NA	3	NA	-0.1208	-1.0262
INT	Data Integration	0.64	NA	3	NA	-0.3448	-0.7373
EQUI	Data Equivocality	0.75	NA	2	NA	1.0148	1.3591
RICH	Data Richness	0.80	NA	5	NA	-1.0436	0.4592
TASKUT	Data Task Utility	0.78	0.81	5	0.68	0.0371	-0.9585
ENACUT	Data Enactment Utility	0.60	0.79	10	0.57	-0.4288	-0.1444
QUALITY	The Quality of DW CR Data for PE	0.74	NA	15	0.70	-0.0815	-0.6632

Figure 5.6 Normality assessment of scales

[Hair Jr. *et al.* 1998, Segars and Grover 1993]. Evidence of convergent validity adds to a researcher's confidence in the construct validity of measures [Schwab 1999].

In Table 5-25, the correlation ranks from .41 to .70 which is considered moderate $p < .001$[Hair Jr. *et al.* 2003]. Better results are offered by the indicators of QUALITY and ENACENH. Convergent validity is exhibited in CUST, COMP, INT, RICH, EQUI (see Table 5-13) and in KNOWENH, and SELFE (see Table 5-20) in factor analysis with loadings rank between .71 and .90 which is considered high and only four fall in the rank .41 to .70 which is considered moderate [Hair Jr. *et al.* 2003]. Being conservative, I can safely claim that convergent validity is moderately exhibited in QUALITY, CUST, COMP, INT, RICH, EQUI, KNOWENH, SELFE, TASKUT and ENACUT.

5.12.4 Generalization of Research Findings Assessment

External generalization should be achieved via replication. Researchers conducting single studies can only speculate about conditions that may apply outside the domain of variables studied. Judgmental claims of generalization, therefore, must be provisional. Whether results actually generalize across potential moderator variables to other cases or contexts can only be established by replication [Schwab 1999]. However, this research includes some elements supporting external validation:

- Lack of moderating effect for the two variables studied JOBFUNC and DWFUNC suggests strong support to the associations in SCT (i.e. there are no exceptions caused by the two variables).
- Qualitative research validated the findings

Given that (i) the representativeness and normality of the sample frame, (ii) the reliability of the sample data and the validity of the research findings, and (iii) the size of the sample all meet the required research standards, it appears reasonable to conclude that the study's findings should be considered for future replication. Only after successful replication of the findings in this study it is possible to claim generalization to the population.

5.13 SUMMARY OF HYPOTHESES NOT SUPPORTED

This chapter has provided a phase-by-phase description of the analyses made executing the research plan. One deviation with respect to the research plan was the additional ANOVA analyses made in phase VI as a result of the small support that the qualitative research gave to the finding that the number of supported CRM data warehouse functions is not a moderator and it does not produce group differences (see Table 5-37).

HYPOTHESES NOT SUPPORTED
H4.3 There is an association between (i) customer relationship problem sense making information search behavioral heuristic patterns HEUR and (ii) customer relationship problem sense making enhancements ENACENH
H4.4 There is an association between (i) customer relationship problem sense making information search behavioral mode patterns MODE and (ii) the quality of data warehouse customer relationship data for problem enactment QUALITY.
H5.1 CRM job function moderates the reciprocal association between (i) the quality of data warehouse customer relationship data for problem enactment, and (ii) customer relationship problem sense-making enhancements.
H5.2 CRM job function produces group differences in the quality of data warehouse customer relationship data for problem enactment.
H5.3 CRM job function produces group differences in customer relationship problem sense-making enhancements.
H5.4 CRM job function produces group differences in customer relationship problem sense-making information search mode.
H5.5 CRM job function produces group differences in customer relationship problem sense-making information search heuristics.
H5.1 The number of supported CRM data warehouse functions moderates the reciprocal association between (i) the quality of data warehouse customer relationship data for problem enactment, and (ii) customer relationship problem sense-making enhancements.
H5.2 The number of supported CRM data warehouse functions produces group differences in the quality of data warehouse customer relationship data for problem enactment.
H5.3 The number of supported CRM data warehouse functions produces group differences in customer relationship problem sense-making enhancements.
H5.4 The number of supported CRM data warehouse functions produces group differences in customer relationship problem sense-making information search mode.
H5.5 The number of supported CRM data warehouse functions produces group differences in customer relationship problem sense-making information search heuristics.

Table 5-37 Hypotheses not supported.

6. RESEARCH FINDINGS

6.1 RESEACH FINDINGS OF PHASE I

Hypothesis H1 was tested in phase I aiming for the discovery of the inner associations of QUALITY.

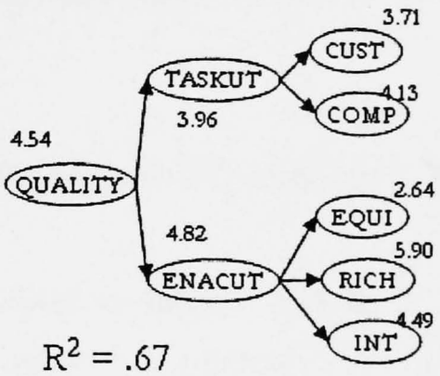


Figure 6.1 Supported hypothesis and research findings on inner associations of QUALITY

Hypothesis 1: The quality of data warehouse customer relationship data for problem enactment QUALITY is a high order latent construct with two dimensions including (i) a context dependent (i.e. CRM) data task utility dimension TASKUT providing instrumental utility for the context's tasks, and (ii) a situation-within-context dependent (i.e. customer relationship problem enactment) data enactment utility dimension ENACUT providing instrumental utility for the considered situation-within-context.

- **Research finding #1**: Hypothesis H1 is supported with R^2 = .67 (see Figure 6.1).

Hypothesis H2 was tested in phase I aiming for the discovery of the inner associations of ENACENH.

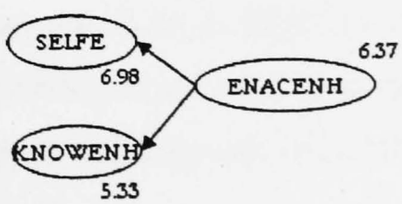


Figure 6.2 Supported hypothesis on inner associations of ENACENH

Hypothesis 2: Customer relationship problem sense making enhancements ENACENH is a high order latent construct with two dimensions including the following two situation-within-context dependent (i.e. customer relationship problem enactment) dimensions: (i) an enactive mastery experience knowledge enhancement dimension KNOWENH reporting self-beliefs on knowledge enhancements in the considered situation-within-context, and (ii) an enactive mastery expectancy self-efficacy dimension SELFE reporting self-efficacy beliefs in prospective situations-within-context.

. **Research finding #2**: Hypothesis H2 is supported with $R^2 = .58$ (see Figure 6.2).

Hypothesis H3.1 was tested in phase I aiming for the discovery of different behavioral patterns in terms of customer relationship problem sense making information search mode orientations.

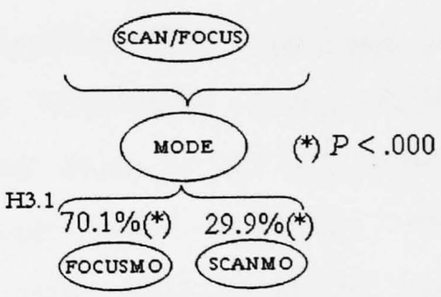


Figure 6.3 FOCUSMO and SCANMO customer relationship problem sense making information search mode orientations

Hypothesis 3.1: There are different behavioral patterns MODE in terms of the exhibited customer relationship problem sense making information search modes SCAN and FOCUS.

. **Research finding #3**: Hypothesis H3.1 is supported (see Figure 6.3). There are two clusters (p <.000) with FOCUSMO mode orientation as the dominant cluster (70.1%), due to its lower mean, $x = 31.52$, in SCAN (i.e. high mean in FOCUS) with repect the second cluster. SCANMO mode orientation (29.9%) has a high mean, $x = 67.67$, in SCAN (i.e. low mean in FOCUS) with respect the first cluster.

Hypothesis H3.2 was tested in phase I aiming for the discovery of different behavioral patterns in terms of customer relationship problem sense making information search heuristic orientation.

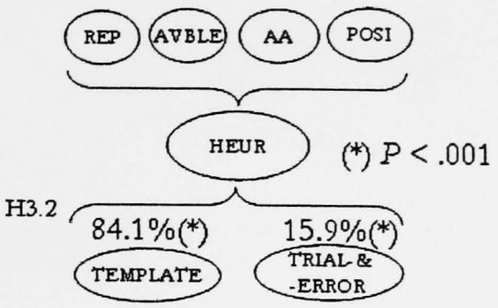


Figure 6.4 Template and trial-and-error customer relationship problem sense making information search heuristic orientations.

Hypothesis 3.2: There are different behavioral patterns HEUR in terms of the exhibited customer relationship problem sense making information search heuristics REP, AVBLE, AA and POSI.

Research finding #4: Hypothesis H3.2 is supported (see Figure 6.4). There are two clusters (p <.001) with TEMPLATE mode orientation as the dominant cluster (84.1%), due to its higher means in REP and AVBLE, x = 25.46 and x = 23.94 respectively, and lower means in AA and POSI, x = 18.62 and x = 17.21 respectively with repect the second cluster. TRIAL-&-ERROR mode orientation (15.9%) is the second cluster with higher means in AA and POSI, x = 42.29 and x = 32.13 respectively, and lower means in REP and AA, x = 15.05 and x = 13.45 respectively with respect the first cluster.

Figure 6.5 and Table 6-1 illustrate the part of the research model that is supported by successfully tested hypotheses in this phase.

6.2 RESEACH FINDINGS OF PHASE II

Hypothesis H4.1 was tested in phase II aiming for the confirmation of the association between QUALITY and ENACENH.

Hypothesis 4.1: There is an association between the quality of data warehouse customer relationship data for problem enactment QUALITY, and customer relationship problem sense making enhancements ENACENH.

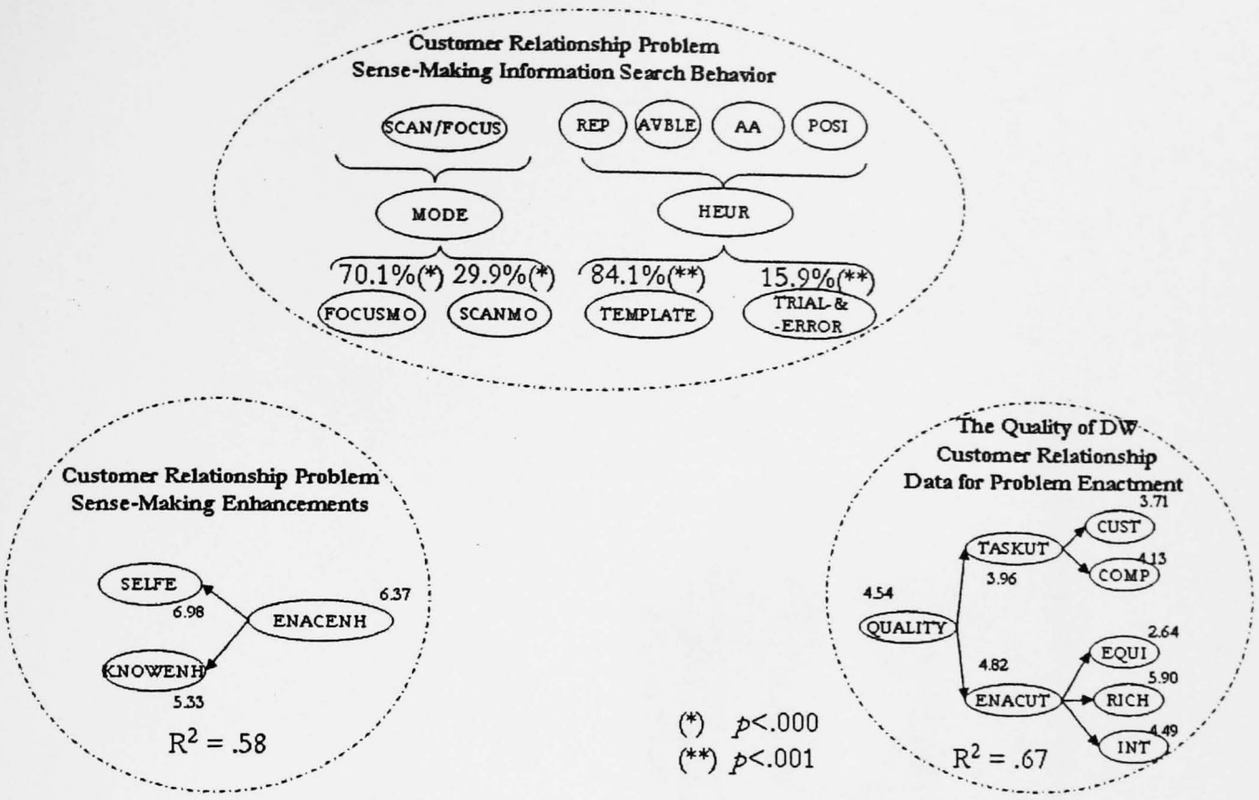


Figure 6.5 Research findings in phase I (means included for QUALITY and ENACENH)

#	RESEARCH FINDINGS Description	(*) p < .000 (**) p < .001 R^2	(***) p = .001 (****) p < .02 %N
1	H1 The quality of data warehouse customer relationship data for problem enactment QUALITY is a high order latent construct with an internal structure including (i) a context dependent (i.e. CRM) data task utility dimension TASKUT providing instrumental utility for the context's tasks, and (ii) a situation-within-context dependent (i.e. customer relationship problem enactment) data enactment utility dimension ENACUT providing instrumental utility for the considered situation-within-context.	67.741	
2	H2 Customer relationship problem sense making enhancements ENACENH is a high order latent construct with an internal structure including the following two situation-within-context dependent (i.e. customer relationship problem enactment) dimensions: (i) an enactive mastery experience knowledge enhancement dimension KNOWENH reporting self-beliefs on knowledge enhancements in the considered situation-within-context, and (ii) an enactive mastery expectancy self-efficacy dimension SELFE reporting self-efficacy beliefs in prospective situations-within-context.	58.376	
3	H3.1 There are different behavioral patterns MODE in terms of the exhibited customer relationship problem sense making information search modes SCAN and FOCUS. FOCUSMO is the predominant customer relationship problem sense-making information search mode orientation.		70.09% (*)
4	H3.2 There are different behavioral patterns HEUR in terms of the exhibited customer relationship problem sense making information search heuristics REP, AVBLE, AA and POSI TEMPLATE is the predominant customer relationship problem sense-making information search heuristic orientation.		86.92% (*) (**)

Table 6-1 Hypotheses supported in phase I

. **Research finding #5**: Hypothesis H4.1 is supported (see Figure 6.7). Furthermore, the association is positive with customer relationship problem sense making enhancements ENACENH as the dependent variable and quality on data warehouse customer relationship data for problem enactment QUALITY as the independent variable The unstandarized coefficient beta indicates that when QUALITY goes up by 1, then ENACENH goes up by .793 (R^2 = .28, p <.000).

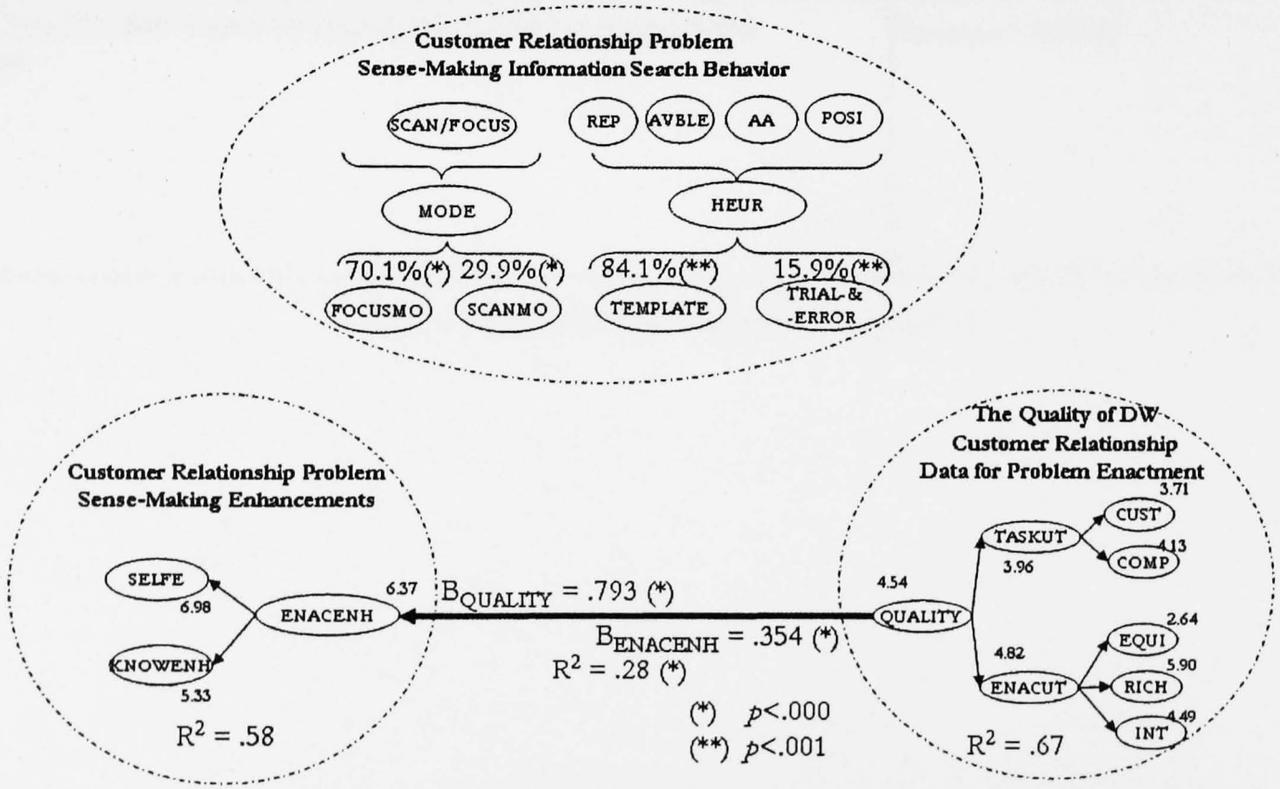


Figure 6.6 Research findings in phases I and II (means included for QUALITY and ENACENH)

RESEARCH FINDINGS		(*) $p < .000$ (**) $p < .001$	(***) $p = .001$ (****) $p < .02$		
#	Description	B	R	SEM Goodness-of-Fit	R^2
5	H4.1There is an association between the quality of data warehouse customer relationship data for problem enactment QUALITY, and customer relationship problem sense making enhancements ENACENH.	$B_{QUALITY}$ = 0.793 (*) $B_{ENACENH}$ = 0.354(*)	0.530 (*)	p of χ^2 GFI RMR standardized RMSEA HOELTER .01 AGFI NFI χ^2/d f	0.28 (*)

Table 6-2 Hypotheses supported in phase II

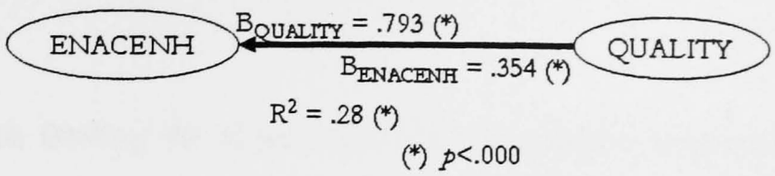


Figure 6.7 A tested recursive model supports the association between QUALITY and ENACENH.

Figure 6.6 and Table 6-2 illustrate the part of the research model that is supported by successfully tested hypotheses in this phase.

6.3 RESEACH FINDINGS OF PHASE III

Hypotheses H4.2, H4.3, H4.4 and H4.5 were tested in phase III aiming for the confirmation of the associations of ENACENH and QUALITY with MODE and HEUR. Hypotheses H4.3 and H4.4 were rejected because discriminant analysis did not find any dependences.

Hypothesis 4.2: There is an association between customer relationship problem sense making information search behavioral mode patterns MODE and customer relationship problem sense making enhancements ENACENH.

Hypothesis 4.5: There is an association between customer relationship problem sense making information search behavioral heuristic patterns HEUR and the quality of data warehouse customer relationship data for problem enactment QUALITY.

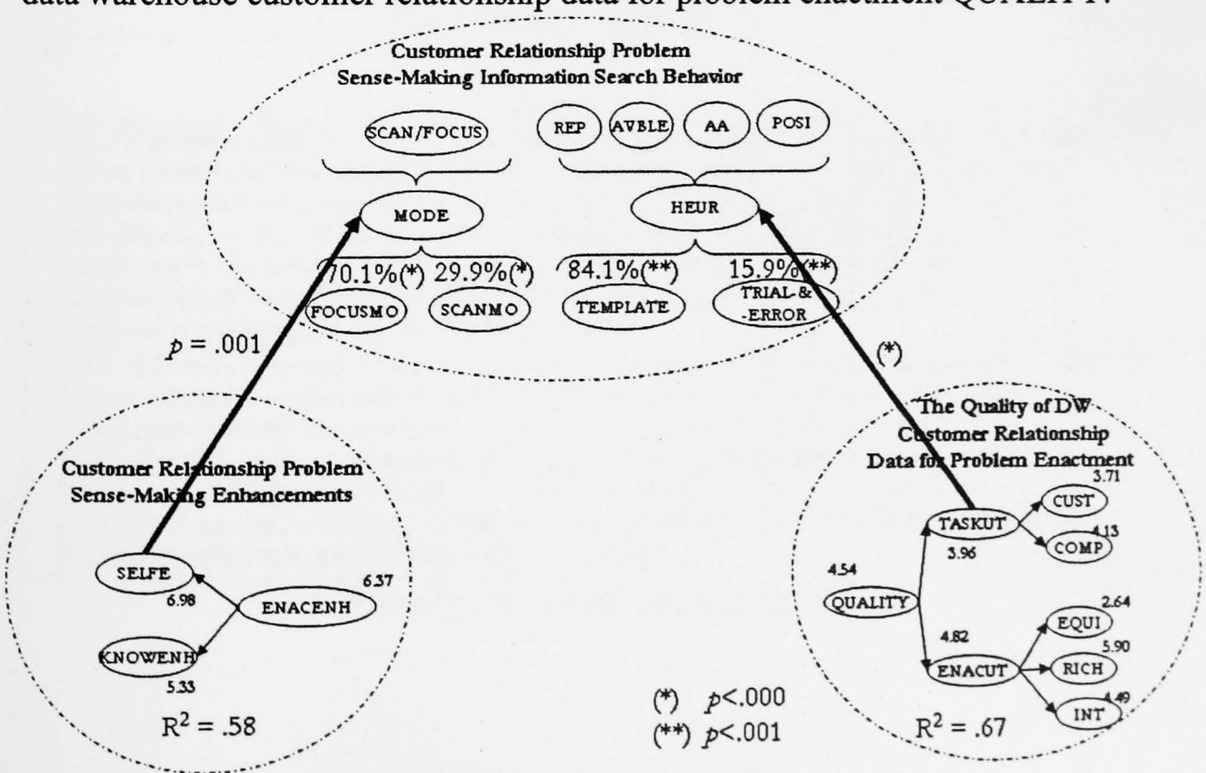


Figure 6.8 Supported associations with behavioral constructs

- **Research finding #6**: Hypotheses H4.2 is partially supported because customer relationship problem enacting self-efficacy SELFE (a dimension in ENACENH) is a discriminant of customer relationship problem sense making information search mode orientations MODE. The high scores in SELFE will lead to a higher likelihood of predicting the SCANMO mode orientation and the low scores in SELFE will lead to a higher likelihood of predicting the FOCUSMO mode orientation (see Figure 6.8).

- **Research finding #7**: Hypotheses H4.5 is partially supported because data task utility TASKUT (a dimension in QUALITY) is a discriminant of customer relationship problem sense making information search heuristic orientations HEUR. The high scores in TASKUT will lead to a higher likelihood of predicting the TEMPLATE heuristic orientation and the low scores in TASKUT will lead to a higher likelihood of predicting the TRIAL-and-ERROR heuristic orientation (see Figure 6.8).

Finally, Table 6-3 shows the research findings in this phase and Figure 6.9 illustrates the part of the research model that, so far, is supported by research findings.

RESEARCH FINDINGS		(*) $p < .000$ (***) $p = .001$
#	Description	Predictive Probability
6	H4.2 is partially supported. There is an association between customer relationship problem sense making information search behavioral mode patterns MODE and customer relationship problem sense making enhancements ENACENH.Customer relationship problem enacting self-efficacy is a discriminant of customer relationship problem sense making information search mode orientations. The high scores in SELFE will lead to a higher likelihood of predicting the SCANMO mode orientation and the low scores in SELFE will lead to a higher likelihood of predicting the FOCUSMO mode orientation	0.656 (***)
7	H4.5 is partially supported. There is an association between customer relationship problem sense making information search behavioral heuristic patterns HEUR and the quality of data warehouse customer relationship data for problem enactment QUALITY. Data task utility is a discriminant of customer relationship problem sense making information search heuristic orientations. The high scores in TASKUT will lead to a higher likelihood of predicting the TEMPLATE heuristic orientation and the low scores in TASKUT will lead to a higher likelihood of predicting the TRIAL-and-ERROR heuristic orientation	.857 (*)

Table 6-3 Research findings in phase III

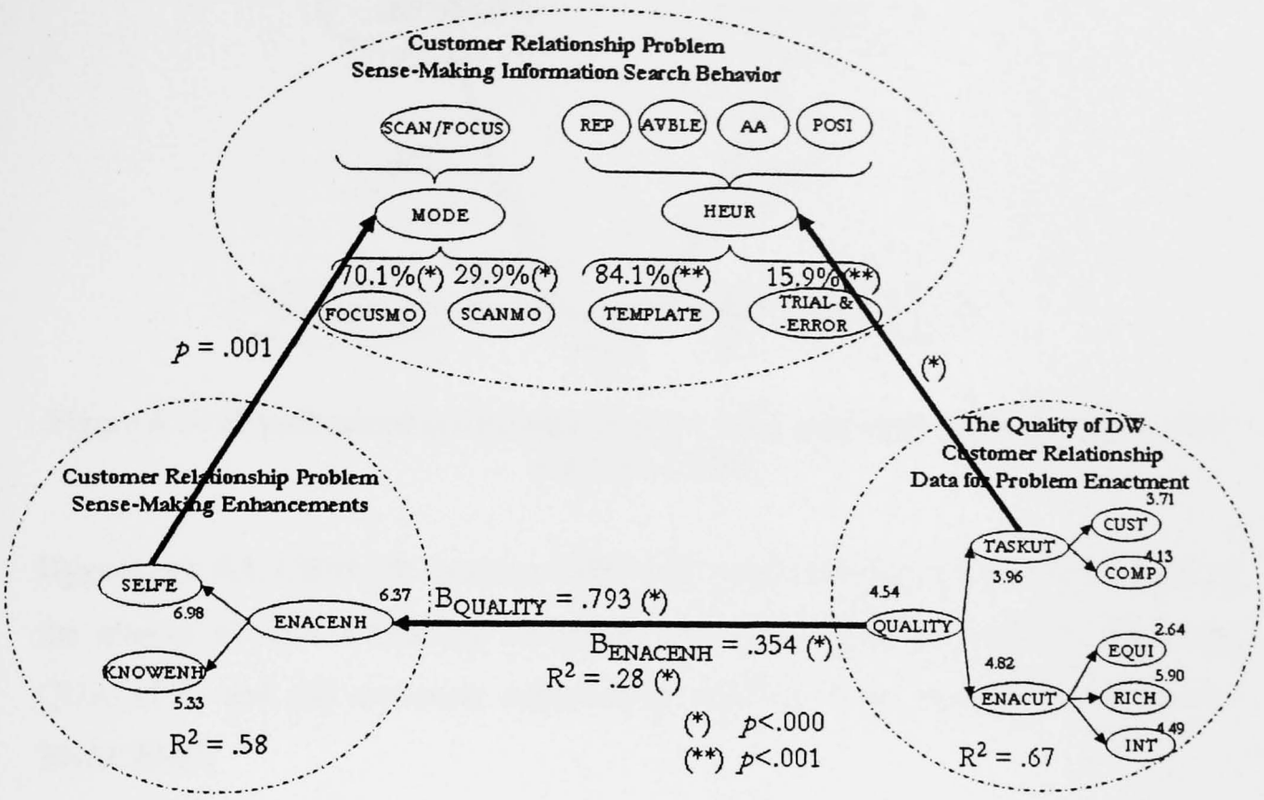


Figure 6.9 Research findings in phases I, II and III (means included for QUALITY and ENACENH)

6.4 RESEACH FINDINGS OF PHASE IV

Hypotheses H5.1 and H6.1 were tested in phase IV aiming for the confirmation of a moderating impact in the association between the quality of data warehouse customer relationship data for problem enactment QUALITY and customer relationship problem sense making enhancements ENACENH. Hypotheses H5.1 (CRM job function moderates this association) and H6.1 (CRM data warehouse function moderates this association) were rejected (see Figure 6.10). This means that CRM job function and the number of supported CRM data warehouse functions are not moderating the association between QUALITY and ENACENH. The hypothesis H6.1 was tested for the two groups considered more relevant (i.e. "just one CRM function supported by the data warehouse" and "more than one CRM function supported by the data warehouse"). Because the two selected groups fully covered all the other groups, it seams reasonable to generalize the finding to all the groups implied by DWFUNC. This lack of support for hypotheses H5.1 and H6.1 is a fascinating finding that is commented on the next chapter.

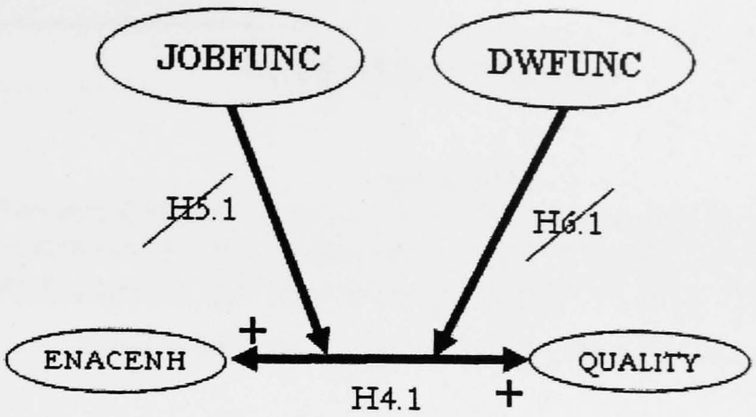


Figure 6.10 Hypothesized moderating impact on the association between QUALITY and ENACENH

Hypothesis 5.1: CRM job function JOBFUNC moderates the association between (i) the quality of data warehouse customer relationship data for problem enactment QUALITY, and (ii) customer relationship problem sense making enhancements ENACENH.

Hypothesis 6.1: The number of suported CRM data warehouse functions DWFUNC moderates the association between (i) the quality of data warehouse customer relationship data for problem enactment QUALITY, and (ii) customer relationship problem sense making enhancements ENACENH.

- **Research finding #8i**: Hypothesis H5.1 is not supported. CRM job function JOBFUNC does not moderate the association between the quality of data warehouse customer relationship data for problem enactment QUALITY and customer relationship problem sense making enhancements ENACENH (see Figure 6.10).

- **Research finding #8ii**: Hypothesis H6.1 is not supported. The number of supported CRM data warehouse functions DWFUNC does not moderate the association between the quality of data warehouse customer relationship data for problem enactment QUALITY, and customer relationship problem sense making enhancements ENACENH (see Figure 6.10).

Finally, Table 6-4 shows the research findings in this phase.

RESEARCH FINDINGS	
#	Description
8	CRM Job Function JOBFUNC is not a moderator. In particular, regardless the CRM job function the studied information search modes are performed with even intensity.
	The number of supported CRM data warehouse functions DWFUNC is not a moderator.

Table 6-4 Research findings supported in phase IV

6.5 RESEACH FINDINGS OF PHASE V

Hypotheses H5.2, H5.3, H5.4, H5.5, H6.2, H6.3, H6.4, and H6.5 were tested in phase IV aiming for the confirmation of group differences in the constructs in the research model produced by CRM job function JOBFUNC (see Figure 6.11) and the number of supported CRM data warehouse functions DWFUNC (see Figure 6.12). All these hypotheses were rejected.

Hypothesis 5.2: CRM job function JOBFUNC produces group differences in the quality of data warehouse customer relationship data for problem enactment QUALITY.

Hypothesis 5.3: CRM job function JOBFUNC produces group differences in customer relationship problem sense making enhancements ENACENH.

Hypothesis 5.4: CRM job function JOBFUNC produces group differences in customer relationship problem sense making information search modes SCAN and FOCUS.

Hypothesis 5.5: CRM job function JOBFUNC produces group differences in customer relationship problem sense making information search heuristics REP, AVBLE, AA, and POSI.

Research finding #9i: Hypotheses H5.2, H5.3, H5.4 and H5.5 are not supported. In particular, the lack of support for the hypothesis H5.4 implies that assumptions that individuals performing the marketing support job function exhibit higher levels of information search scanning mode orientation than individuals performing other CRM functions are not supported by this research (see Figure 6.11).

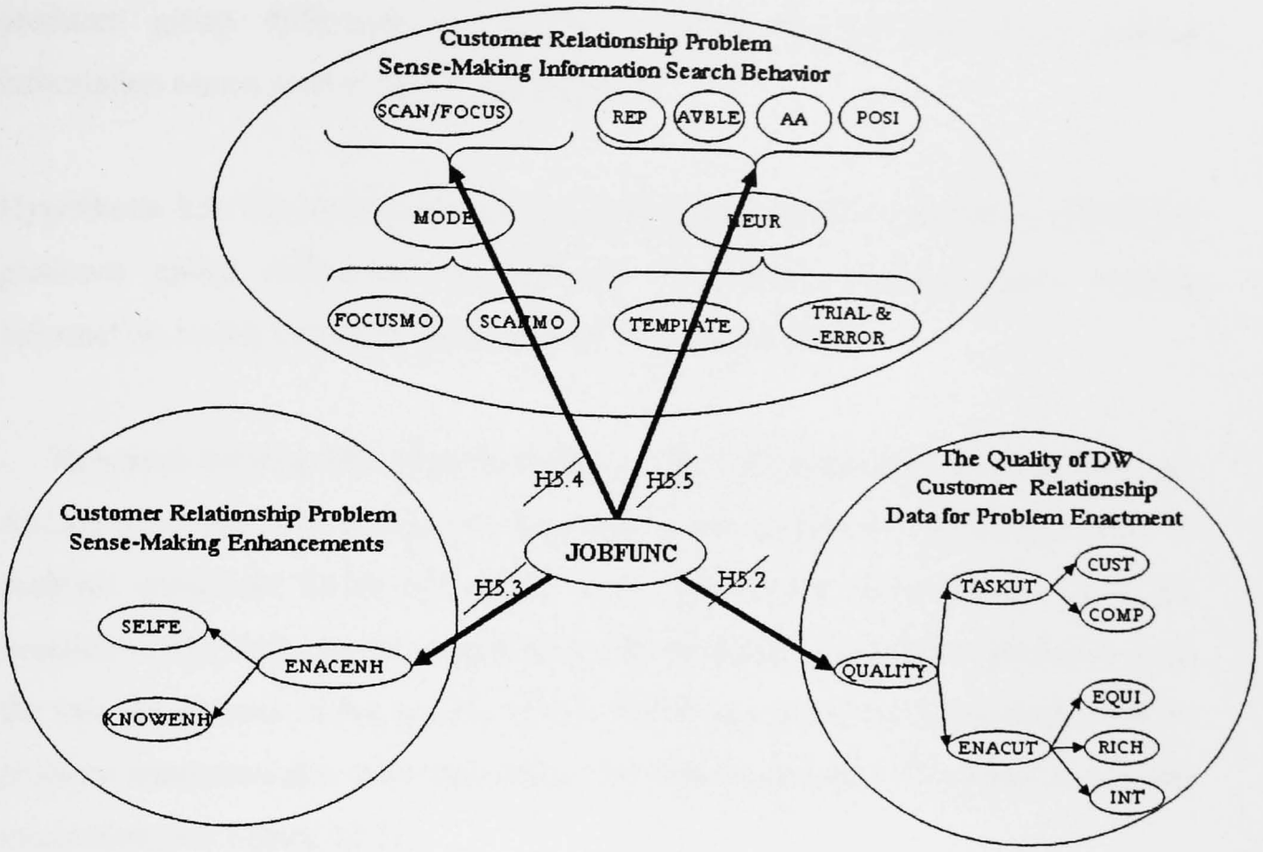


Figure 6.11 Hypothesized group differences produced by JOBFUNC

Likewise, the number of supported CRM data warehouse functions DWFUNC is not producing any group differences in the quality of data warehouse customer relationship data for problem enactment QUALITY, customer relationship problem sense making enhancements ENACENH, customer relationship problem sense making information search modes SCAN/FOCUS, or customer relationship problem sense making information search heuristics REP, AVBLE, AA, POSI.

Hypothesis 6.2: The number of supported CRM data warehouse functions **DWFUNC** produces group differences in the quality of data warehouse customer relationship data for problem enactment QUALITY.

Hypothesis 6.3: The number of supported CRM data warehouse functions DWFUNC produces group differences in customer relationship problem sense making enhancements ENACENH.

Hypothesis 6.4: The number of supported CRM data warehouse functions DWFUNC produces group differences in customer relationship problem sense making information search modes SCAN and FOCUS.

Hypothesis 6.5: The number of supported CRM data warehouse functions DWFUNC produces group differences in customer relationship problem sense making information search heuristics REP, AVBLE, AA, and POSI.

. **Research finding #9ii**: Hypotheses H6.2, H6.3, H6.4 and H6.5 are not supported. Assumptions that the quality of data warehouse customer relationship data for problem enactment QUALITY grows with the number of supported CRM job functions DWFUNC are not supported by this research. This is also applicable to (i) the two dimensions of the quality of data warehouse customer relationship data for problem enactment (i.e. data task utility and data enactment utility), and to (ii) data integration (see Figure 6.12).

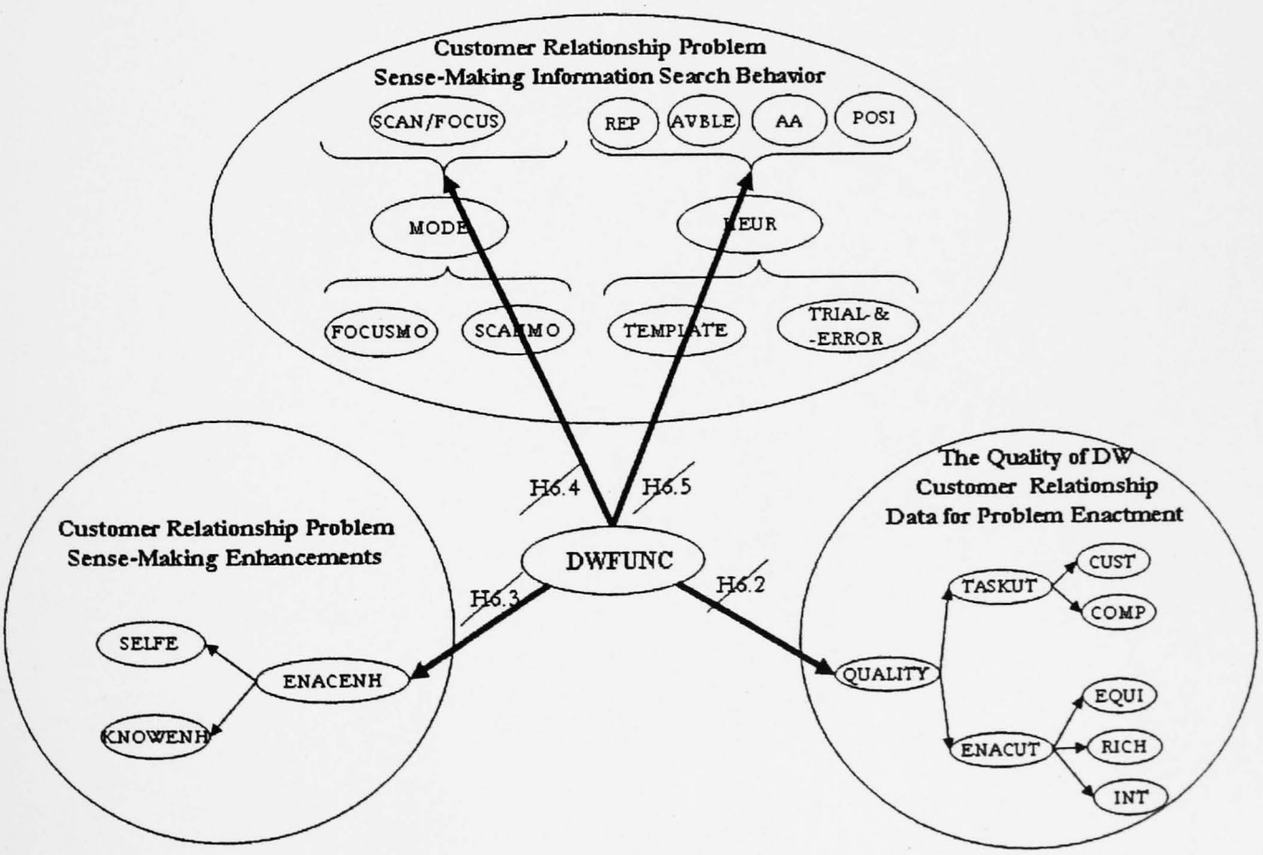


Figure 6.12 Hypothesized group differences produced by DWFUNC

Finally, Table 6-5 shows the research findings in this phase.

RESEARCH FINDINGS	
#	Description
9	Hypotheses H5.2, H5.3, H5.4 and H5.5 are not supported. In particular, the lack of support to the hypothesis H5.4 implies that assumptions that individuals performing the marketing support job function exhibit higher levels of information search scanning mode orientation than individuals performing other CRM functions are not supported by this research.
	Hypotheses H6.2, H6.3, H6.4 and H6.5 are not supported. Assumptions that the quality of data warehouse customer relationship data for problem enactment QUALITY grows with the number of supported CRM job functions DWFUNC are not supported by this research. This is also applicable to (i) the two dimensions of the quality of data warehouse customer relationship data for problem enactment (i.e. data task utility and data enactment utility), and to (ii) data integration.

Table 6-5 Research findings supported in phase V

6.6 RESEACH FINDINGS OF PHASE VI

Qualitative findings in this phase are summarized in Table 6-6.

	RESEARCH FINDINGS	INTERVIEWEES' LEVEL OF SUPPORT	INTERPRETATION AS A RESULT OF OPEN QUESTIONS
1	Internal structure of QUALITY (H1)	Strongly agree	Direct benefits: Tracking end-users perceptions, Indicator for corrective action Indirect benefits: Image of taking into account the opinion of internal customers
	QUALITY scores	Moderately agree	Surprise (e.g. expectations of lower scores in QUALITY, expectations of higher scores in INT, expectations that CUST scores would be higher than COMP scores).
2	Internal structure of ENACENH (H2)	Strongly agree	Direct benefits: Tracking end-users self-perceptions
	ENACENH scores	Strongly agree	No surprise, meets experience.
3	There are different situation-whithin-context customer relationship problem sense-making information search mode orientations (H3.1). They are (i) focus orientation and (ii) scanning orientation.	Strongly agree	No surprise, meets expectations, disappointment.
	Focus is the predominant information search mode orientation		
4	There are different situation-whithin-context customer relationship problem sense-making information search heuristic orientations (H3.2). They are (i) template orientation and (ii) trial-and-error orientation.	Strongly agree	No surprise, meets expectations, disappointment.
	Template is the predominant information search heuristic orientation		
5&6&7	QUALITY is an indepedent variable of ENACENH	Strongly agree	No surprise, meets expectations. Opportunities for improvement.
	SELFE is a discriminant of customer relationship problem sense making information search mode orientations FOCUSMO vs. SCANMO.	NA	NA
	TASKUT is a discriminant of customer relationship problem sense making information search heuristic orientations TEMPLATE vs. TRAIL-and-ERROR.	NA	NA
	Overall high-level model explaining the supported associations between environmental, cognitive and behavioral constructs	Strongly agree	It makes sense.
8&9	CRM job function JOBFUNC is not a moderator and it does not produce group differences	Strongly agree	No surprise, unfortunately meets expectations (e.g. very disappointing evidence that marketing people predominant information search mode is focus instead of scanning).
	The number of supported CRM data warehouse functions DWFUNC is not a moderator and it does not produce group differences	Slightly agree	Surprise. It conflicts with expectations that the more CRM supported functions by the DW, the higher should be the integration and therefore the higher the overall quality scores.

Table 6-6 Qualitative findings validating and interpreting quantitative findings

7. CONCLUSIONS

7.1 INTRODUCTION

The chapter starts with a section summarizing the main contributions of this research. These contributions are further explained in separate sections, each one providing a partial answer to the research question. The section WHAT IS THE NATURE OF THE QUALITY OF DATA WAREHOUSE CUSTOMER RELATIONSHIP DATA FOR PROBLEM ENACTMENT? presents the deconstruction of the data quality construct. The literature suggests that enhancements in problem sense making are a consequence of data quality. Unfortunately, few studies address this association. The same happens with respect to the association between data quality and information search behavior. I argue that this situation has favored a lack of conceptual clarity in relation to the nature of both enhancements in problem sense making and information search behavior constructs. Therefore, a needed step in understanding such constructs was to understand their nature.

The section WHAT IS THE NATURE OF CUSTOMER RELATIONSHIP PROBLEM SENSE MAKING ENHANCEMENTS? covers conclusions addressing the deconstruction of enhancements in problem sense making in the context of this research (i.e. customer relationship management supported by a data warehouse) while the section WHAT IS THE NATURE OF CUSTOMER RELATIONSHIP PROBLEM SENSE MAKING INFORMATION SEARCH BEHAVIOR? covers conclusions addressing the deconstruction of information search behavior, again, in the context of this research (i.e. customer relationship management supported by a data warehouse). The sections IS DATA QUALITY? and IS DATA QUALITY THE DEPENDENT VARIABLE? cover conclusions addressing the associations with data quality as the independent and dependent variable respectively. The section WHAT IS THE IMPACT OF THIRD VARIABLES? summarizes the conclusions on the impact of the variable 'CRM job function' and the variable 'number of supported CRM data warehouse functions' as moderators and producers of group differences in the research model.

The section ARE COGNITIVE ENHANCEMENTS ASSOCIATED TO ? presents conclusions about the associations involving data quality. This approach is consistent with SCT in relation to the existence of triadic associations between environmental, cognitive and behavioral constructs. The RESEARCH LIMITATIONS AND WEAKNESSES section covers what can be done better and the RESEARCH AGENDA section covers suggestions for the next research steps. Finally, as a result of the findings, a number of guidelines have been produced and they are presented in the RECOMMENDATIONS FOR A DATA QUALITY PRACTICE section.

7.2 KEY CONTRIBUTIONS

My research question is *What are the inner and inter construct associations of the quality of data warehouse customer relationship data for problem enactment*? The overall contribution of this research is to give clarity to the nature of data quality in an integrated framework covering technical, cognitive and behavioral aspects. This general statement has five significant contributions:

(i) The literature on data quality has neglected the specific inner associations of this construct. This research contributes to the body of knowledge with the specification of the inner associations of data quality. Two dimensions were identified, a contextual dimension and a situational dimension. This scale is new in the literature.

(ii) The literature on information systems benefits and sense making has systematically neglected the inner associations of cognitive enhancements. This research contributes to the body of knowledge with the specification of the inner associations of problem sense making enhancements. Two dimensions were identified, a knowledge enhancement dimension and a self-efficacy dimension. This scale is new in the literature.

(iii) Information search behavioral patterns are not frequently studied in the literature. This research contributes to the body of knowledge with the specification of behavioral patterns for information search mode (scanning mode orientation vs.

focus mode orientation) and for information search heuristic (template heuristic orientation vs. trial-and-error heuristic orientation).

(iv) A comprehensive confirmatory study of SCT is a challenge of considerable proportions. To the best of my knowledge I only know a handful of comprehensive studies addressing Bandura's SCT as a whole (i.e. most of the studies cover them in a partial way) and generally they are done in a controlled environment. This research has contributed to the body of knowledge by testing SCT in a comprehensive way and using an empirical approach. In particular, this research has provided empirical evidence on the following:

- Enhancements in problem enactment and data quality are associated. Furthermore, the research results provide strong evidence supporting data quality as the independent variable and enhancements in problem enactment as the dependent variable.
- Data task utility (the contextual dimension of data quality) and information search heuristic orientation are associated. Data task utility (the contextual dimension of data quality) is a discriminant variable of the information search heuristic orientations template versus trial-and-error.
- Self-efficacy (dimension of enhancements) and information search mode orientation are associated. Self-efficacy is a discriminant variable of the information search mode orientations focus versus scanning.
- The variable 'CRM job function' and the variable 'number of supported CRM data warehouse functions' have no impact as moderators and/or producers of group differences.

This research contributes to the body of knowledge with the specification of the data quality for problem enactment model, DQ4PEM (see Figure 7.4), and the data quality cognitive metamodel (see Figure 7.5) DQ[C, SwC] that illustrates a family of data quality models in a context (e.g. CRM, ERP) and a situation in such context (e.g. enacting problems, solution building). Then, the DQ4PEM would be a subset of the data quality cognitive metamodel for problem enactment situations (i.e. DQ4PEM = DQ[C, SwC] with C= {CRM, ERP,...} and SwC= Problem enacting).

7.3 WHAT IS THE NATURE OF THE QUALITY OF DATA WAREHOUSE CUSTOMER RELATIONSHIP DATA FOR PROBLEM ENACTMENT?

Data quality has been studied in this research in the context of CRM supported by a data warehouse and in a type of situation in such context characterized by the enactment of problems (i.e. the formulation of a problem statement). The construct 'quality of data warehouse customer relationship data for problem enactment' is defined as *the manager's perception on the extent to which the information derived from the available customer relationship data in the data warehouse fits with the manager's customer relationship management cognitive needs for problem enactment.* Individuals develop a cognitive structure that includes types of information when they formulate a problem statement. The 'fit' term, in the definition of data quality, means that there is a match between the types of required information in a developed cognitive structure and the types of available information. Perfect quality means perfect match, and therefore, the problem statement could be formulated with zero uncertainty. On the other hand, the worse possible quality means no match at all and therefore, the uncertainty is absolute for the task of enacting a problem. Data quality defined in the former terms is the reverse construct of uncertainty enacting customer relationship problems.

The research findings suggest that the nature of the quality of data warehouse customer relationship data for problem enactment is a high order construct with two dimensions defined in terms of (i) a context (CRM in this research) and (ii) a situation-within-context, problem enactment in this research. These inner associations confirm the theory (e.g. [Brannick 2000]). Also, the research findings suggest that another situation in the same context should require adaptation of the data enactment utility dimension. However, another context would require both an adaptation of the data task utility dimension for the new context and an adaptation of the situation-within-context dimension to the new situation. Both dimensions are second order constructs. The contextual dimension, data task utility, has in its turn two dimensions, one for customer insights and the other for competitor insights. The situational-within-context dimension, data enactment utility, has in its turn three dimensions, one

measuring data integration, one for data richness and another for data equivocality (see Figure 7.1).

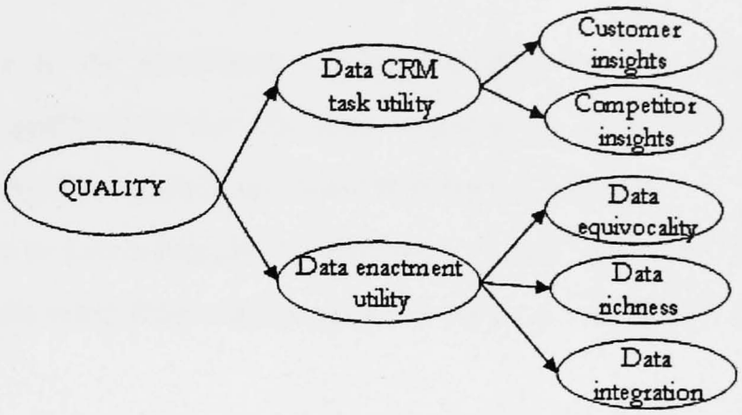


Figure 7.1 The Quality of DW Customer Relationship Data for Problem Enactment Inner Associations (QUALITY)

In simple terms, the emerged nature of data quality in this research suggests that data quality is specific to a situation and a context and has a dimension for each of them. This implies that generalistic approaches to data quality are of limited help because they are not focused. Another implication is that replication of studies about data quality should be very thorough and report the context, situation, and unit of analysis. Otherwise, comparisons about scores might make little sense.

The scores in each of the components in the iner associations of the quality of data warehouse customer relationship data for problem enactment have the following interpretations:

- In relation to the quality of data warehouse customer relationship data for problem enactment, informants slightly agree that the information derived from the available customer relationship data in the data warehouse fits with their needs when they have to enact customer relationship problems.
- In relation to the contextual dimension, CRM data task utility, informants neither agree nor disagree that the data warehouse customer relationship data has instrumental utility for CRM tasks.
- In relation to customer insights, informants neither agree nor disagree that the data warehouse's customer relationship data is a source of customer insights for CRM tasks.

- In relation to competitor insights, informants slightly agree that the data warehouse's customer relationship data is a source of competitor insights for CRM tasks.
- In relation to the situational-within-context dimension, data enactment utility, informants slightly agree that the data warehouse's customer relationship data has enactment utility for enacting customer relationship problems.
- In relation to data integration, informants slightly agree that the data warehouse's customer relationship data is normalized in terms of data definitions and logical data structures.
- In relation to data richness, informants moderately agree that the data warehouse's customer relationship data is a source of customer relationship information.
- In relation to data equivocality, informants slightly disagree that the data warehouse's customer relationship data favors more than one interpretation for the enactment of customer relationship problems.

Interestingly, the customer insights mean is slightly lower than the competitor insights mean and when considering the standard deviation, both measures practically overlap. This conflicts with the expectation of the usual availability of both types of data, being that the former is more available than the latter. One observation to be made here is that the wording of the scales made clear that lack of availability means bad quality.

Another remark is the data integration mean. One would expect that talking about data warehouses and with 95 out of 107 cases supporting more than one CRM function, data integration would have a higher mean than a modest "slightly agree". Likewise, informants do not seem to be very enthusiastic with the data task utility dimension giving it just a "neither agree nor disagree." Overall, these research findings show that there is room for improvement in data quality in data warehouses supporting CRM processes.

7.4 WHAT IS THE NATURE OF CUSTOMER RELATIONSHIP PROBLEM SENSE MAKING ENHANCEMENTS?

Enhancements in problem sense making have been studied in this research in the context of customer relationship management supported by a data warehouse and in a type of situation in such context characterized by the enactment of problems (i.e. the formulation of a problem statement). The construct 'customer relationship problem sense making enhancements' is defined as *the manager's self-belief about the extent that the enactment of customer relationship problems has improved due to the data warehouse's customer relationship data.*

The research findings suggest that the nature of problem sense making enhancements is a high order construct with two dimensions both defined in terms of a situation-within-context. In this research the situation-within-context is the enactment of customer relationship problems in the context of customer relationship management supported by a data warehouse. One dimension looks backwards, that is, the extent that one's skills enacting customer relationship problems have improved due to the data warehouse customer relationship data. The other dimension looks forward, that is the strength in the self-belief in one's skills to execute given types of performances enacting prospective customer relationship problems. These dimensions confirm the theory (e.g. [Bandura 1997]). As with the data quality construct, a different research situation should require adaptation of both dimensions in this case as opposed to just one like with the data quality construct (see Figure 7.2).

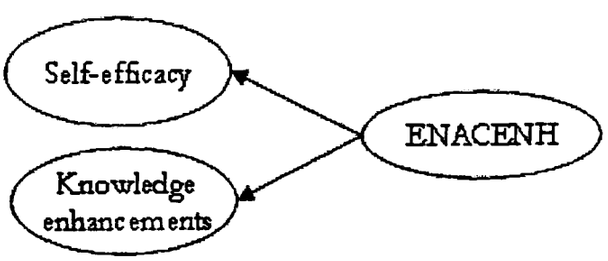


Figure 7.2 Customer Relationship Problem Sense-Making Enhancements Internal Inner Associations (ENACENH)

The scores in each of the components in the inner associations of customer relationship problem sense making enhancements have the following interpretations:

- In relation to the customer relationship problem sense making enhancements, informants strongly agree that their enactment of customer relationship problems has improved due to the data warehouse's customer relationship data.
- In relation to the customer relationship problem enactment knowledge enhancement, informants moderately agree that their skills enacting customer relationship problems have improved due to the data warehouse's customer relationship data.
- In relation to the customer relationship problem enacting self-efficacy, informants strongly agree that their skills will allow them to execute given types of performances enacting prospective customer relationship problems due to the data warehouse's customer relationship data.

Interestingly, informants seem to be very confident in relation to their future skills because the customer relationship problem enacting self-efficacy mean is clearly higher than the customer relationship problem enactment knowledge enhancement mean. A temptation would be to make conclusions of over-confidence, balanced confidence and under-confidence based on the trend of customer relationship problem enactment knowledge enhancement vs. customer relationship problem enacting self-efficacy. These types of assessments have been well studied in the decision confidence literature (e.g. [Chung 1995, Aldag and Power 1986, Benbasat and Schroeder 1977]) but they might not be appropriate for enacting problems. Overall, the research findings clearly indicate benefits attributed to the data warehouse customer relationship data in informants' ability enacting customer relationship problems.

7.5 WHAT IS THE NATURE OF CUSTOMER RELATIONSHIP PROBLEM SENSE MAKING INFORMATION SEARCH BEHAVIOR?

The research findings suggest that there are two types of information search behavior that contribute to explain the nature of customer relationship problem sense making information search behavior. Both types of behavior are defined in terms of the situation-within-context and by the manager's perceptions on the personal's amount of effort using them. In this research the situation-within-context is the enactment of

customer relationship problems. One type of information search behavior is the search mode and the other is the search heuristic.

The construct 'customer relationship problem sense making information search mode' is defined as *the manager's information search behavior characterized by the type of inquiry.* The information search modes considered in this research are intensity of data warehouse scanning search mode and intensity of data warehouse focused search mode. Scanning search is the proactive and exploratory information search behavior mode people exhibit when they browse through information without a particular problem to solve. Focused search is the reactive and directed information search behavior mode people exhibit when they are looking for information specific to a problem to be addressed or question to be answered.

The construct 'customer relationship problem sense making information search heuristics' is defined as *the information search behaviour characterized by the type of a simplifying routine.* The information search heuristics considered in this research are availability search heuristic, representativeness search heuristic, anchoring and adjustment search heuristic, and positivy search heuristic. Availability search heuristic refers to assessing the probability of a situation as a function of prior situations (e.g. a marketer considering a series of occurrences of actual costs incurred in past editions of a campaign when estimating the cost for a new edition of such campaign). This heuristic implies the search for easily accessible information about relevant precedents. The search ends once recent, salient information about a relevant precedent is found. Representativeness search heuristic refers to (i) assessing the probability of a situation as a representative of a category (e.g. a 'price-lowering by a competitor' situation can have common information with an 'attempt action to gain market-share' pattern), or to (ii) making generalizations based on new information about a sample - i.e. the sample is representative of a large population (e.g. to assess the national market success of a new product line based on the data likelihood ratio of a test market). This heuristic implies that the search ends when there is a satisficing fit between information about a situation and information about a category.

Anchoring and adjustment search heuristic refers essentially to the trial and error method (e.g. a marketer setting the price of a product starting with a baseline price and making a number of impact analyses in several of the cost components). This heuristic implies a recursive process and each step involves a search for additional information and an adjustment of the previous assessment. The search ends once the adjustments are not improving an implicit/explicit value function on the information found. Positivy search heuristic refers to confirming the probability of a situation using the trial and error method (e.g. a marketer looking for issues in accounts that were predicted to have a high risk). This heuristic implies a search for information that is fundamentally consistent with existing beliefs, theories and cognition. The search ends once the information found confirms the probability of a situation.

The research findings suggest that there are two patterns in search mode and two other patterns in search heuristics (see Figure 7.3).

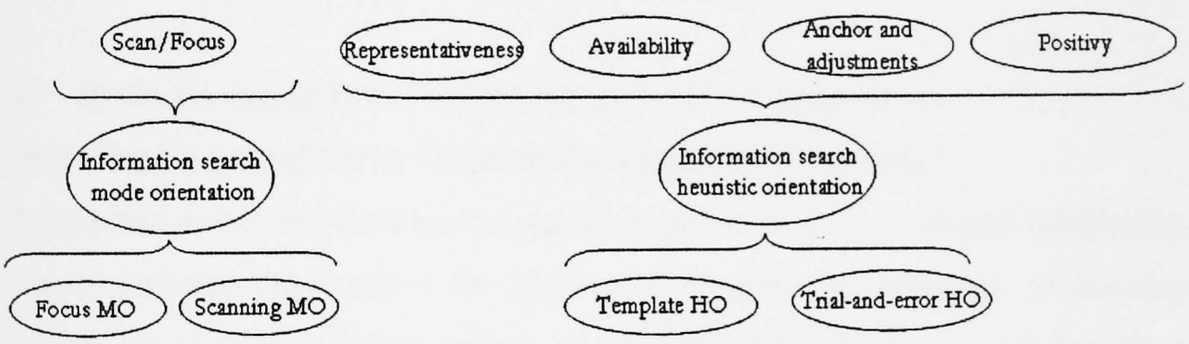


Figure 7.3 Customer Relationship Problem Sense-Making Information Search Behavior Inner Associations

Individuals concurrently exhibit a pattern of search mode and a pattern of search heuristic. In the search mode behavior, there are two patterns. One labeled as scanning mode orientation, that is, an emphasis on scanning search mode over focused search mode. The other labeled focus mode orientation, that is, an emphasis on focused search mode over scanning search mode. In this research, the predominant pattern is focus mode orientation (75 cases out of 107).

In the search heuristic behavior, there are two patterns. One labeled as template heuristic orientation, that is, an emphasis on representativeness and availability search

heuristics over anchor and adjustment and positivy search heuristics. Template heuristic orientation involves an information template either as the reference while searching representative information or as the goal of the search. The other search behavior pattern was labeled as trial-and-error heuristic orientation, that is, an emphasis on anchor and adjustment and positivy search heuristics over representativeness and availability search heuristics. This orientation follows the trial-and-error approach. In this research, the predominant pattern is TEMPLATE heuristic orientation (93 cases out of 107).

In simple terms, this means that a CRM manager is likely to search for information following a focus mode orientation and a template heuristic orientation. This overall conclusion is consistent with the garbage can model (e.g. [Cohen, March and Olsen 1972]) where the templates take the role of 'solutions'. Later, this overall conclusion will be qualified with comments about the null role of CRM job function as a factor introducing group differences.

7.6 IS DATA QUALITY THE INDEPENDENT VARIABLE?

7.6.1 Are Cognitive Enhancements the Dependent Variable?

The positive association between the quality of data warehouse customer relationship data for problem enactment as the independent variable and customer relationship problem sense making enhancements as the dependent variable means that (in a positive sense) improvements in data quality will be associated to enhancements in the ability of CRM managers' formulating problem statements. Furthermore, it is amazing that this association is between a modest "slightly agree" level on data quality and a "strongly agree" level on problem sense making enhancements. In simple terms, a mediocre level on data quality is associated to a big difference in enacting CRM problems.

The research findings in this study are based in associational analysis and not in causal analysis. Therefore, they cannot be framed in causation terms of "consequences". However, they provide strong evidence that the quality of data warehouse customer relationship data for problem enactment is the independent variable of customer relationship problem sense making enhancements which

supports SCT (e.g. [Bandura 1997]), Vessey's theory of cognitive fit (e.g. [Vessey 1991]) and Simon's human information processing theory (e.g. [Simon 1976]). Their theoretical propositions postulate that the environment (e.g. data quality in this study) influences cognitive factors (e.g. problem enactment in this study). For example, the situational dimension found in this study in data quality (i.e. problem enactment utility) has data integration as one of its components. Data integration results in greater task accuracy and faster task completion [Goodhue *et al.* 2000]. This implies that enhancements in the task of enacting problems with respect to the same task spanning over dispersed (i.e. non-integrated) information will be higher because non-integrated data will favor more errors as more complex mental processing will be required (e.g. [Goodhue *et al.* 2000]).

Another example, the contextual dimension found in this study in data quality (i.e. CRM data task utility) is organized around customer data and competitor. Data organization may provide better or worse information for the task at hand [Newell and Simon 1972]. This implies that enhancements in the task of enacting problems with well organized CRM data respect the same task using poorly organized data will be higher because poorly organized data will provide worse information (e.g. [Seddon 1997]).

Altough the evidence found in this study is not sufficient for claiming that causation has been proven this finding provides empirical evidence for an argument not yet empirically tested, to the best of my knowledge: The higher the data quality, the higher the cognitive enhancements.

7.6.2 Are Behavioral Information Search Patterns the Dependent Variable?

SCT (e.g. [Bandura 1997]) and adaptive behavior theories (e.g. [Payne 1976]) postulate that environmental conditions (e.g. data quality in this study) influence behavior (e.g. information search heuristics in this study). Although no single heuristic does well across all situations, a person can maintain a reasonably high level of accuracy at a low level of effort by selecting from a repertoire of strategies contingent upon situational demands (e.g. [Payne, Bettman and Johnson 1993, Beach

and Mitchell 1978]). This implies that the level of data quality influences the selection of information search heuristics.

Although the research findings of this study cannot be framed in causation terms of "consequences" they are consistent with these theories. The research findings of this study provide evidence that the contextual dimension of data quality, CRM data task utility, is a discriminant variable of information search heuristic orientations template versus trial-and-error. For example, higher scores in CRM data task utility will be associated to a higher likelihood of predicting the template heuristic orientation and the lower scores in CRM data task utility will be associated to a higher likelihood of predicting the trial-and-error heuristic orientation. In simple terms, CRM managers exhibiting a template information search heuristic orientation operate with data warehouses with higher levels on CRM data task utility than the data warehouses with which CRM managers operate exhibiting a trial-and-error information search heuristic orientation. Again, this overall conclusion will be qualified later with a conclusion about the null role of number of CRM functions supported by the data warehouse as a factor introducing group differences. Surprisingly, according to the research findings no quality construct is playing a similar role in information search mode orientation, that is discriminating between the scanning and focus mode orientations.

7.7 IS DATA QUALITY THE DEPENDENT VARIABLE?

7.7.1 Are Cognitive Enhancements the Independent Variable?

The research findings provide little support to the proposition that cognitive enhancements is the independent variable of perceptions on data quality. In this direction (i.e. cognitive enhancements as the independent variable of perceptions on data quality) the strength of the association is roughly 50% less than in the other direction, and this association is between a "strongly agree" level on problem sense making enhancements and a "slightly agree" level on data quality.

7.7.2 Are Behavioral Information Search Patterns the Independent Variable?

This research did not find evidence that information search behavioral patterns are the independent variable of perceptions on data quality.

7.8 WHAT IS THE IMPACT OF THIRD VARIABLES?

The impact of the variable 'CRM job function' and the variable 'number of supported CRM data warehouse functions' as moderators and producers of group differences was studied in this research. The variable 'CRM job function' reports the CRM job function primarily performed by the unit of analysis in this research (i.e. CRM manager). Four CRM job functions were studied: Sales support, marketing support, customer services support and other CRM function. The variable 'number of supported CRM data warehouse functions' reports the CRM functions supported by the data warehouse. Four CRM functions were studied: Sales support, marketing support, customer services support and other CRM function. This last variable had a multiple response format. I considered that the most interesting groups were the group that I labeled as "just one CRM function supported by the data warehouse" and the group that I labeled as "more than one CRM function supported by the data warehouse". It should be understood that when I refer to the variable 'number of supported CRM data warehouse functions' it is meant the selected two groups.

This research found that 'CRM job function' is not moderating, against expectations, the association between data quality and problem sense making enhancements. This finding implies that the scales of data quality and problem sense making enhancements have commonalities to the CRM jobs performed by informants.

Likewise and against expectations, this research found that the 'number of supported CRM data warehouse functions' is not moderating the association between data quality and problem sense making enhancements. This finding implies that the number of CRM functions supported by a data warehouse does not matter when looking at, for example, the cognitive benefits of data quality. This is a very conflictive finding because it contradicts a common argument in the industry that a data warehouse supporting more company functions, the more the better, it enables a

better understanding of customers. This finding was challenged by the participants in the research findings focus group.

When looking at the impact of 'CRM job function' as a producer of group differences in data quality, problem sense making enhancements, and information search behavior (mode and heuristic orientations), this research found, against expectations, that 'CRM job function' is not producing any group difference in any of these constructs. In particular, this finding implies that no CRM job function group is exhibiting a predominant information search scanning mode orientation. The expectation was that informants performing the marketing support function would exhibit such behavioral pattern. However, according to this research, the answer to the question as to which CRM function (i.e. marketing vs. sales vs. etc.) is mainly scanning (i.e. looking for trends, and understanding patterns) is no one because this pattern is evenly exhibited by all the groups. This was acknowledged as a disappointing confirmation of a 'suspect' in the qualitative research (phase VI).

Likewise, looking at the impact of the 'number of supported CRM data warehouse functions' as a producer of group differences in data quality, problem sense making enhancements, and information search behavior (mode and heuristic orientations), this research found, also against expectations, that the 'number of supported CRM data warehouse functions' is not producing any group difference in any of such constructs. In particular, this finding implies that the number of CRM functions supported by a data warehouse when looking at, for example, the scores of data quality does not matter. This is another very conflictive finding because it contradicts a common argument in the industry claiming that a data warehouse supporting more company functions, the more the better, implies higher levels of data quality. This finding was challenged by the participants in the research findings focus group. A potential explanation might be that the notion of a data warehouse supporting several CRM functions does not imply high data integration (i.e. a dimension in data quality). In simple terms, you could have a data warehouse supporting many CRM functions and very low scores in the data integration dimension. A potential subsequent challenge derived from this conclusion could be the following: Is data quality a valid measure for comparing two data warehouses in the same situation-within-context

studied in this research, one supporting just one CRM function and the other supporting several CRM functions? If the answer is yes, why? The answer is yes. Furthermore, it will be very informative for benchmarking purposes. The justification for the affirmative response is that (i) the inner associations of data quality, supported by the findings in this research, considers data integration as a relevant quality dimension addressing more than one CRM functions and (ii) 'number of supported CRM data warehouse functions' does not produce group differences according to the findings in this research.

7.9 ARE COGNITIVE ENHANCEMENTS ASSOCIATED TO BEHAVIORAL INFORMATION SEARCH PATTERNS?

Bandura's SCT [1986] and Simon's human processing information theory (e.g. [Simon 1976]) postulate that cognitive factors (e.g. problem enactment in this study) control behavior (e.g. information search mode in this study). From a means-ends rationality perspective the means are contingent with respect to the ends (e.g. [Simon 1977]). Means-ends rationality implies that the 'means' (e.g. information search mode in this study) should be appropriate to reach the desired ends (e.g. problem enactment is this study). Therefore, the selection of an information search mode is influenced by problem enactment. From a SCT perspective, the stronger the perceived self-efficacy, the higher the goals people set for themselves [Wood and Bandura 1989]. Goals attached to scanning search mode are considered more challenging than the goals attached to focused search mode (e.g. [Shaver and Scott 1991]). This implies that scanning search mode is associated with stronger self-efficacy than focus search mode.

Although the research findings of this study cannot be framed in causation terms of "consequences" they are consistent with these theories. The research findings of this study provide evidence that the self-efficacy dimension of cognitive enhancements is a discriminant variable of information search mode orientations focus versus scanning. For example, higher scores in self-efficacy will lead to a higher likelihood of predicting the scanning mode orientation and the lower scores in self-efficacy will lead to a higher likelihood of predicting the focus mode orientation. In simple terms,

CRM managers exhibiting a scanning information search mode orientation have higher levels of confidence of their ability in enacting prospective CRM problems than CRM managers exhibiting a focus information search mode. Therefore, this research has found that information search mode orientations are associated with self-efficacy which has practical relevance.

7.10 RESEARCH LIMITATIONS AND WEAKNESSES

7.10.1 Research Limitations

A test of the detailed model using structural equation modeling would have required a sample size N=345 for a full confirmatory factor analysis. This size would also have had the benefit of bigger groups in the analysis of the impact effect of 'CRM job function' and 'number of supported CRM data warehouse functions'. Unfortunately, the sample size in this research is N=107, therefore I had to apply a reductionist approach considering the first order latent constructs as observed variables. I argue that this reductionist approach has influenced in the low construct validity of QUALITY. Therefore, the results of the analysis with SEM are not conclusive.

Normalized (to 100) summated scales introduced a multicollinearity effect on the "other/last" option. I dropped this "other/last" heuristic from my analysis. This is a design limitation of these types of scales that I am formally acknowledging here but of difficult solution. If the "other" function is dropped, this would imply conceptual completeness of the rest of the items supported by the existing literature. This was not the case in the scales addressing the information search behavior. Additionally, informants would be forced to fit their answer to a response format that might produce sample error. On the other hand, if you introduce the "other" option it automatically becomes a dependent variable of the rest of the options introducing a multicollinearity issue. I consciously decided to minimize sample error and accept this limitation.

A dummy variable was created based on the fifteen possible combinations that the multiple response format of the 'number of supported CRM data warehouse functions' variable allows. This dummy variable covered the group "just one CRM function supported by the data warehouse" and the group labeled "more than one

CRM function supported by the data warehouse". This introduces a serious limitation. The 'number of supported CRM data warehouse functions' is not a producer of group differences for the two groups studied, but this cannot "automatically" be extended to the fifteen possible combinations. The only reason for this limitation is research effort and time analyzing all the combinations.

Discriminant analyses and ANOVA analyses require the same size of clusters/groups. This means that the sample size for such analyses had to be adjusted in order to meet such a requirement. The sample size for research finding #7 was N=64, for research finding #8 was N=28, for research findings #9i and #10i was N=82, and for research findings #9ii and #10ii was N=24.

Although the questionnaire was administered only in English, this research has the usual limitation of cultural factors (e.g. [Zhu 1999, Asheghian and Ebrahimi 1990]) in cross-country studies [Straub 1989]. Cultural factors are part of the environment and they are a variable impacting research from a dual perspective: Impacting the researcher (e.g. [Sinickas 2005]) and the observations (e.g. [Sinickas 2005]). This is exacerbated in cross-cultural studies. For example, in relation to the types of constructs that this research have addressed, there are studies showing cross-cultural differences in cognitive structures (e.g. [Bagranoff, Houghthon and Hronsky 1994]), patterns of searching for on-line information (e.g. [Morahan-Martin 2004], use of information technology (e.g. [Johns, Smith and Strand 2002]), decision making (e.g. [Mann, Radford, Burnett, Ford, Bond, Leung, Nakamura, Vaughan and Yang 1998]), response style (e.g. [Chen, Lee and Stevenson 1995]) and response biases (e.g. [Keillor, Owens and Pettijohn 2001]).

In order to mitigate this impact I followed the guidelines in [Sinickas 2005]:

- Checking for interpretation of questions: I conducted a Content Validation Focus group, and a Face validation focus group. The participants in such focus groups were a representative sample of the population. The survey was administered only in English. English is a common communication language among the targeted informants (i.e. market researchers) that do not have English as their native language.
- Pre-testing on-line survey questionnaires with a pilot study

7.10.2 Research Weaknesses

An intrinsic weakness of research based in self-reported input is the validity of the input. Past research based on subjective judgment is endless. Although there is extensive literature using the constructs that are addressed in this study, it is prudent to recognize this weakness.

Additionally, responses were filtered in order to avoid input coming from informants that they fall in the type A data quality (e.g. [Brannick 1998]) that is is ignorance on the quality level that is needed for the task at hand. However, it is not possible to have certainty that this has not happened.

The limitation of the sample size is also a weakness for the accuracy of the confirmatory factor analysis because my reductionist approach (i.e. considering the first order latent constructs as observed variables) introduced some error in the measurement model.

The intensity of the data warehouse anchoring and adjustment search heuristic had departures from normality. This has no effect on cluster analysis but it is controversial in discriminant and ANOVA analyses [Hair Jr. *et al.* 1998]. However: (i) my discriminant analysis did not consider anchoring and adjustment search heuristic as the dependent variable but a cluster with the four heuristics – anchoring and adjustment search heuristic included. Either way, anchoring and adjustment search heuristic for the randomly selected cases in the discriminant analysis (i.e. in order to have same cluster size) did not have normality issues in anchoring and adjustment search heuristic (Skewness= 0.929, Kurtosis=0.216). In relation to the ANOVA analyses, hypotheses H5.5 and H6.5 should be cautiously interpreted with respect to the anchor and adjustment heuristic because of the normality issue (i.e there are no comments with respect to the other heuristics). Overall, although I formally recognize a normality issue in the anchoring and adjustment search heuristic, it had no serious impact in my research findings.

Although the reliability of the situational-within-context quality dimension, data enactment utility, is acceptable for exploratory research (0.60), in terms of relative comparison to the other constructs, it is the least reliable.

7.11 RESEARCH AGENDA

A full confirmation of SCT (e.g. [Bandura 1997]) requires confirmation of the triadic associations among cognitive, environmental and behavioral variables. The findings of this research have successfully confirmed associations in the three pairs of combinations of variables of the research model. The supported hypotheses of this research suggest a contextually generic research model, labeled data quality for problem enactment model, DQ4PEM (see Figure 7.4), which postulates that such triadic association exists. DQ4PEM offers a research framework for situations that require the enactment of problems. DQ4PEM and the scales should be adapted to the specifics of the studied context.

According to the cognitive fit theory (e.g. [Vessey 1991]), individuals develop a cognitive structure that includes types of information when they want to formulate a problem statement. Data quality for problem enactment is defined in DQ4PEM in terms of the cognitive fit theory, as the manager's perception on the extent to which the information derived from the available data fits with the manager's cognitive structures for problem enactment. The 'fit' term, in the definition of data quality, means that there is a match between the types of required information in a developed cognitive structure and the types of available information.

In DQ4PEM, data quality is a high order construct with two dimensions defined in terms of (i) a contextual dimension labeled data task utility and (ii) a situation-within-context dimension labeled data enactment utility. Data task utility is defined as the manager's perception on the extent to which data has instrumental utility for the task at hand (e.g. manufacturing, marketing). Data enactment utility is defined as the manager's perception on the extent to which data has utility for the formulation of problem statements. The data enactment utility dimension has in its turn three dimensions: one measuring data integration, one for data richness and another for data equivocality. In simple terms, the nature of data quality in DQ4PEM suggests

that data quality is specific to a situation and a context and has a dimension for each of them.

The construct 'problem sense making enhancements' is defined as the manager's self-belief about the extent that the enactment of problems has improved due to the available data. In DQ4PEM, problem sense making enhancements is a high order construct with two dimensions. One dimension looks backwards, that is, the extent that one's skills enacting customer relationship problems have improved due to the data. The other dimension looks forward, that is the strength in the self-belief in one's skills to execute given types of performances enacting prospective customer relationship problems.

The DQ4PEM suggests that there are two types of information search behaviors that contribute to explain the nature of problem sense making information search behavior. Both types of behavior are defined in terms of the situation-within-context and by the manager's perceptions on the personal's amount of effort using them. In DQ4PEM the situation-within-context is the enactment of problems. One type of information search behavior is the search mode and the other is the search heuristic. The construct 'problem sense making information search mode' is defined as the manager's information search behavior characterized by the type of inquiry (e.g. [Churchman 1971]).

The information search modes considered in DQ4PEM are intensity of scanning search mode and intensity of focused search mode. Scanning search is the proactive and exploratory information search behavior mode people exhibit when they browse through information without a particular problem to solve (e.g. [Aguilar 1967]). Focused search is the reactive and directed information search behavior mode people exhibit when they are looking for information specific to a problem to be addressed or question to be answered (e.g. [Huber 1991]).

The construct 'problem sense making information search heuristics' is defined as the information search behavior characterized by the type of a simplifying routine. The information search heuristics considered in this research are availability search

heuristic, representativeness search heuristic, anchoring and adjustment search heuristic, and positivy search heuristic. Availability search heuristic (e.g. [Tversky and Kahneman 1972]) refers to assessing the probability of a situation as a function of prior situations (e.g. a marketer considering a series of occurrences of actual costs incurred in past editions of a campaign when estimating the cost for a new edition of such campaign).

This heuristic implies the search for easily accessible information about relevant precedents. The search ends once recent, salient information about a relevant precedent is found. Representativeness search heuristic (e.g. [Kahneman and Tversky 1972]) refers to (i) assessing the probability of a situation as a representative of a category (e.g. a 'price-lowering by a competitor' situation can have common information with an 'attempt action to gain market-share' pattern), or to (ii) making generalizations based on new information about a sample -i.e. the sample is representative of a large population (e.g. to assess the national market success of a new product line based on the data likelihood ratio of a test market). This heuristic implies that the search ends when there is a satisficing fit between information about a situation and information about a category. Anchoring and adjustment search heuristic (e.g. [Tversky and Kahneman 1974]) refers essentially to the trial and error method (e.g. a marketer setting the price of a product starting with a baseline price and making a number of impact analyses in several of the cost components).

This heuristic implies a recursive process and each step involves a search for additional information and an adjustment of the previous assessment. The search ends once the adjustments are not improving an implicit/explicit value function on the information found. Positivy search heuristic (e.g. [Evans 1989]) refers to confirming the probability of a situation using the trial and error method (e.g. a marketer looking for issues in accounts that were predicted to have a high risk). This heuristic implies a search for information that is fundamentally consistent with existing beliefs, theories and cognition. The search ends once the information found confirms the probability of a situation.

The DQ4PEM suggests that there are different patterns in (i) search mode, (ii) search heuristics, and that (iii) individuals concurrently exhibit a pattern of search mode and a pattern of search heuristic. With respect to the links amount the referred constructs, the DQ4PEM suggests association among data quality, problem sense making enhancements and information search behaviors (mode and heuristic orientations).

Furthermore, this research suggests a data quality cognitive metamodel (see Figure 7.5) DQ[C, SwC] that illustrates a family of data quality for a given context (e.g. CRM, ERP) and situation in such context (e.g. enacting problems, solution building). Then, the DQ4PEM would be a subset of the data quality cognitive metamodel for problem enactment situations (i.e. DQ4PEM = DQ[C, SwC] with C= {CRM, ERP,...}and SwC= Problem enacting).

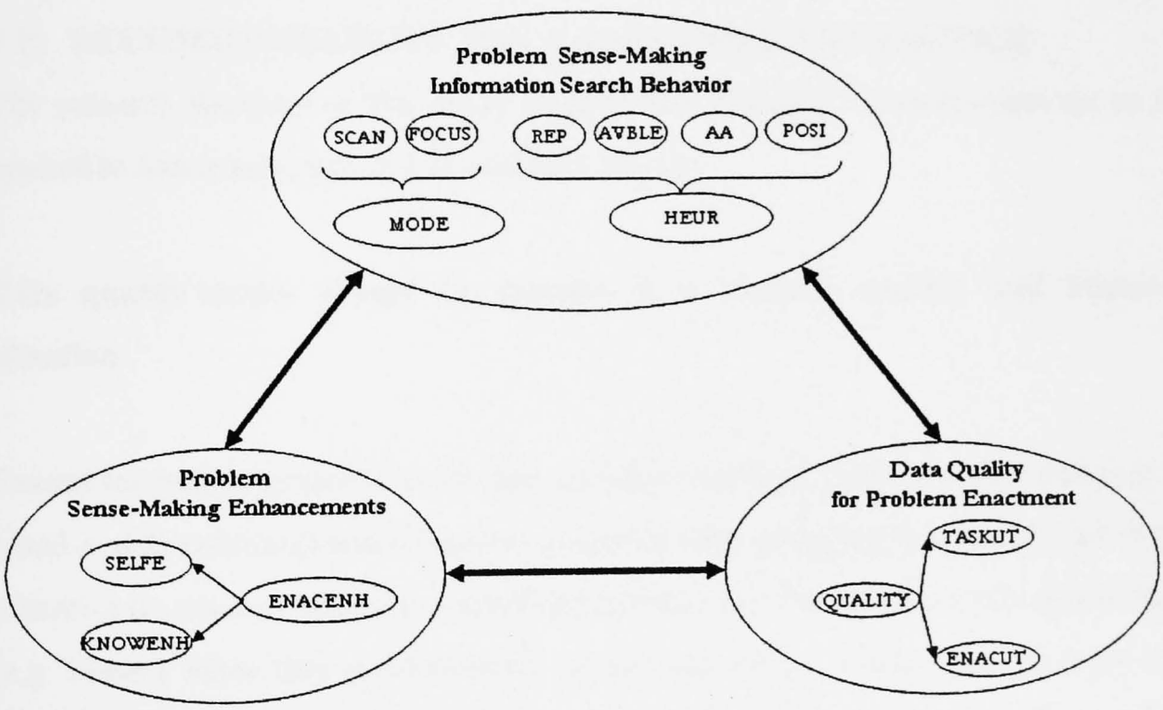


Figure 7.4 Data quality for problem enactment model (DQ4PEM)

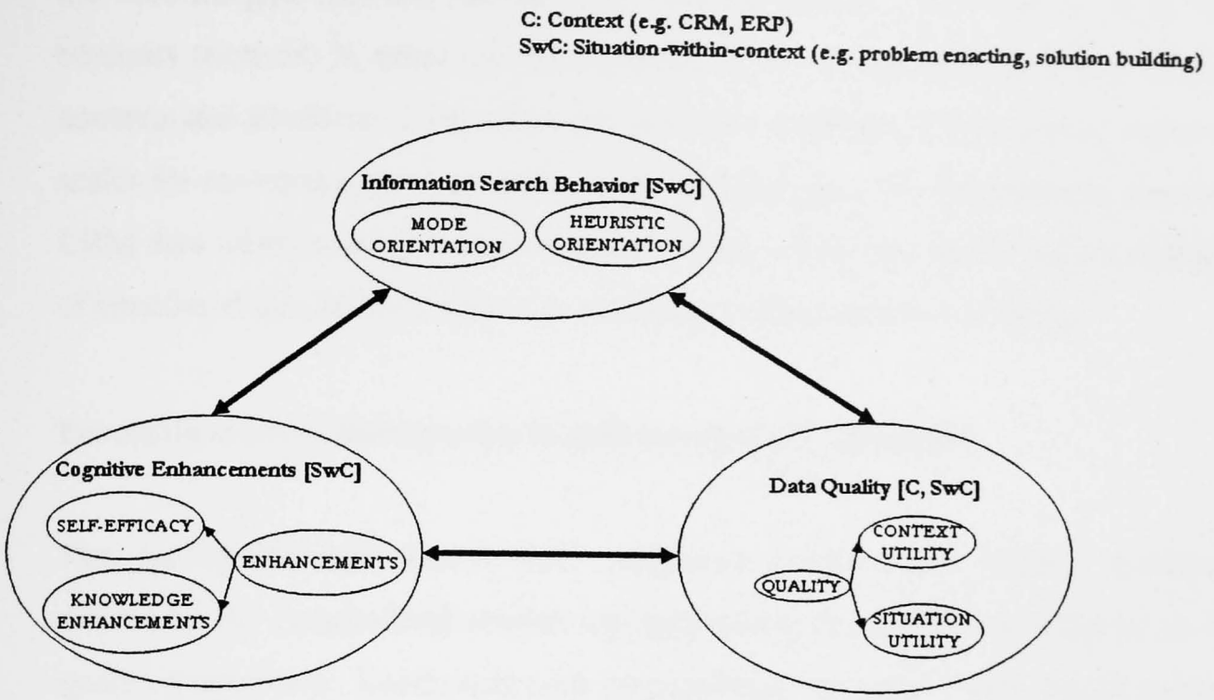


Figure 7.5 Data quality cognitive metamodel DQ[C, SwC]

7.12 RECOMMENDATIONS FOR A DATA QUALITY PRACTICE

The research findings of this study suggest the following recommendations to be applied to data quality practice in business settings.

Data quality scales should be specific to a business context and business situation

Executives with responsibility for data quality should use both objective scales (e.g. based in data profiling) and subjective scales for measuring data quality. In case those measures do not converge, the marketing literature and the decision making literature (e.g. biases) show that user/customer perceptions have influence on attitudes and behavior. This study has provided evidence about the high order nature of data quality in terms of a contextual dimension and a situational dimension. Generic scales (i.e. any context, any situation) for measuring data quality are not recommended because their lack of specificity in the intended utility of data will necessarily result in measurement error. The business context will influence the types of data that the tasks at hand need in order to be performed. For example, product/part data will be centric in an enterprise resource planning context while it might be almost neglected in a customer relationship management context. Therefore, it seems reasonable that

the data insights that are relevant in a business context might not be so in other business contexts. In summary, data quality scales will be different across business contexts and situations. Executives should have a repertoire of data quality situational scales for covering a given context as indicated in Figure 7.6. For example, together a CRM data utility scale (i.e. the contextual scale), executives should have a repertoire of situational data utility scales (e.g. problem enacting, solution building).

Perceptions about data quality should be regularly measured

The quality literature shows that continuous improvement requires continuous measurement. Longitudinal studies are appropriate for measuring variation in data quality perceptions. Executives with responsibility for data quality should measure data quality regularly. Analyses should be done not just to the highest level but to all the levels in the internal structure of data quality in order to gain deep understanding on where the variation is.

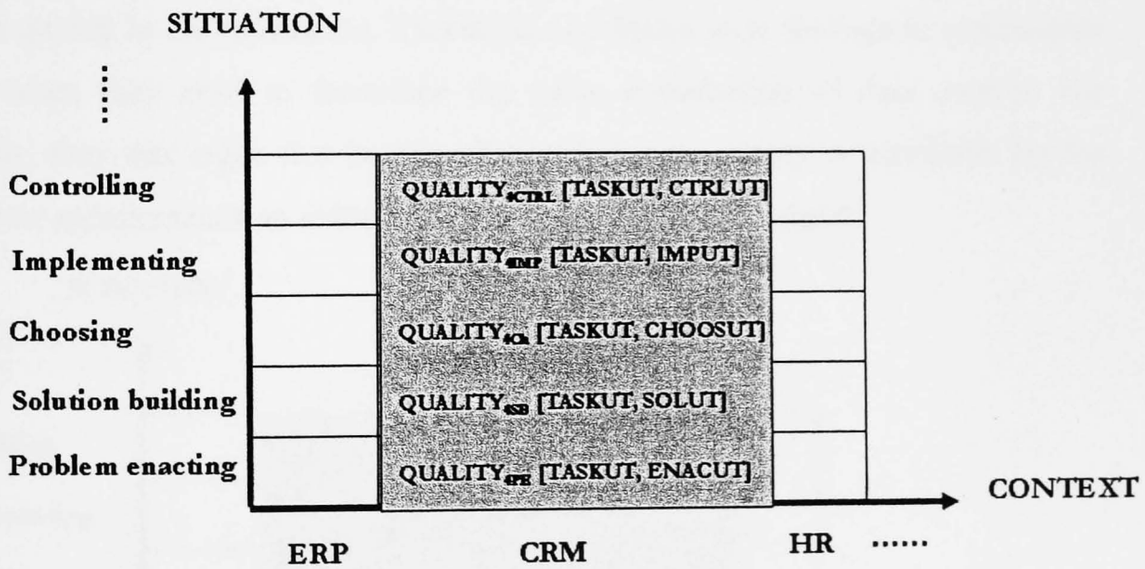


Figure 7.6 Example of a repertoire of data quality situational scales in a CRM context

Enhancement scales should be specific to a business context and business situation

Executives with responsibility for data quality should use both objective scales (e.g. based on return on investment) and subjective scales for measuring the impact of data quality. This study has provided evidence about the high order nature of perceptions of enhancements experienced by end users attributed to data quality in terms of a

knowledge dimension and a self-efficacy dimension. Again, generic scales (i.e. any context, any situation) for measuring enhancements are not recommended because of their lack of specificity in the task and situation will necessarily result in measurement error. Executives should have a repertoire of enhancement situational scales (e.g. problem enacting, solution building) for covering a given business context as indicated in Figure 7.7

Perceptions about enhancements should be regularly measured

Executives with responsibility for data quality should measure end users' enhancements attributed to data quality regularly. Again, analyses should be done not just to the highest level but to the knowledge enhancement and self-efficacy levels in order to gain deep understanding on where the variation is. Executives should study the correlation of enhancements with data quality on longitudinal bases. This is relevant, because executives can demonstrate a positive impact, as this study shows, of data quality in enhancements. Executives can frame such findings in means-ends terms when they need to formulate the value contribution of data quality. For example, they can argue that budget allocated to data quality is justifiable by the impact on enhancements in controlling situations in a CRM context.

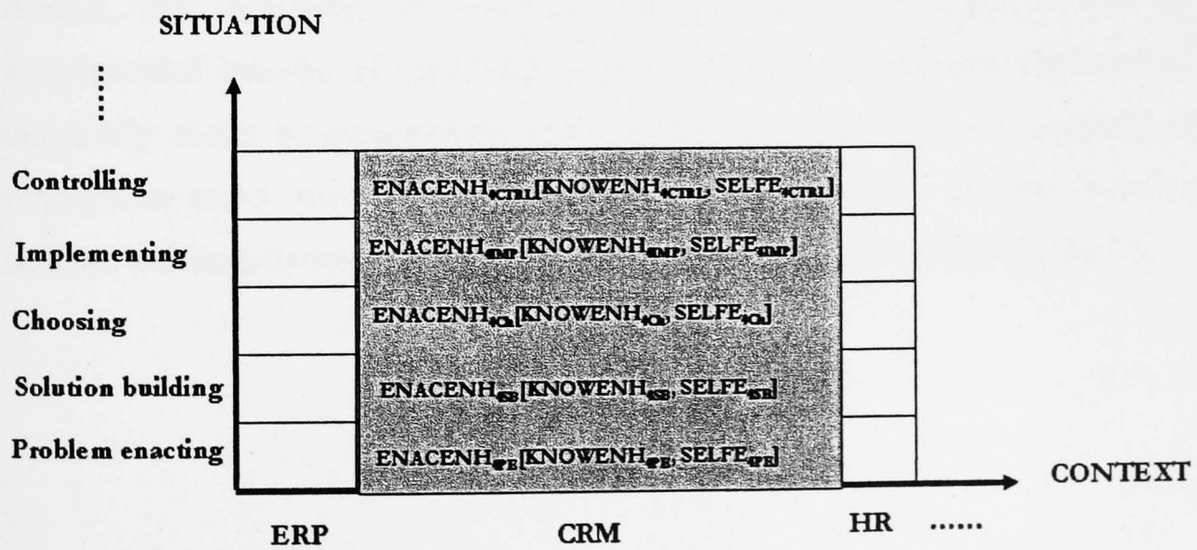


Figure 7.7 Example of a repertoire of enhancement situational scales in a CRM context

Information search behavior scales should be specific to a business context and business situation

Executives with responsibility for data quality should use both objective scales (e.g. based in computer logs) and subjective scales for measuring information search behavior. Effectiveness and efficiency in information search behavior are key considerations that information systems executives should monitor. For example, from a mode orientation perspective, executives should understand if efficiency in information search (i.e. obtain the desired data minimizing the consumption of resources and/or time) is more relevant in a focus orientation than in a scanning orientation. Another example, from a heuristic orientation perspective, executives should understand if higher levels of effectiveness in information search (i.e. obtain the desired data) is more relevant in a trial-and-error orientation than in a template orientation. Such findings would influence recommendations about the technological infrastructure that users should use as a function of their intended information search mode orientation and heuristic orientation.

This study has provided evidence about the nature of information search behavior in terms of a mode orientation and a heuristic orientation. Again, generic scales (i.e. any context, any situation) for measuring information search behavior are not recommended because of their lack of specificity in the task and situation will necessarily result in measurement error. Executives should have a repertoire of information search mode and heuristic situational scales (e.g. problem enacting, solution building) for covering a given business context as indicated in Figure 7.8.

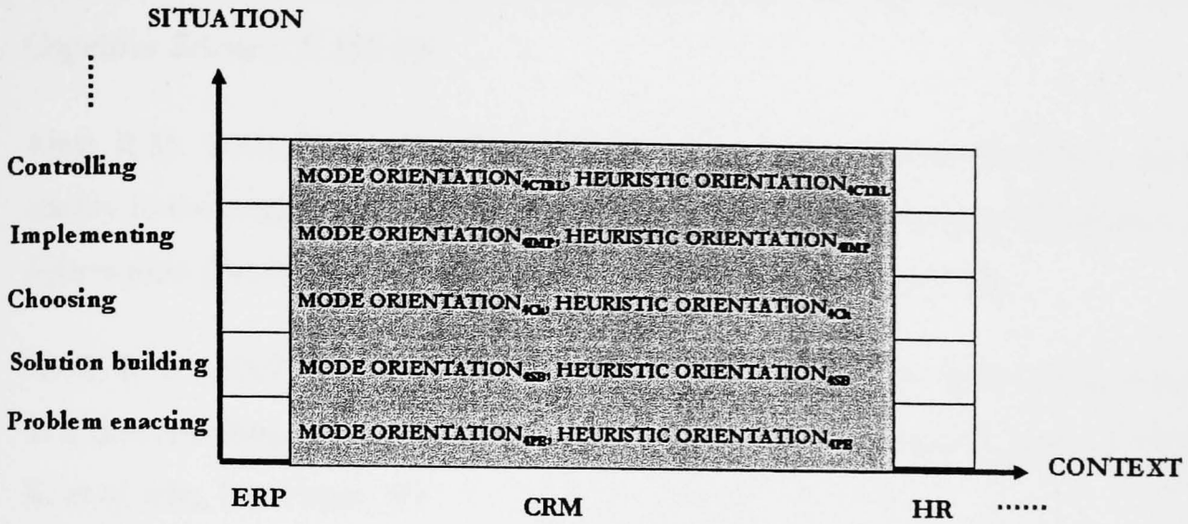


Figure 7.8 Example of a repertoire of information search behavior situational scales in a CRM context

Information search behavior should be regularly measured

Executives with responsibility for data quality should measure information search behavior regularly. Again, analyses should be done for the exhibited mode and heuristic orientations in order to gain deep understanding on where the variations are. Executives should study the correlation of information search behavior with data quality on longitudinal bases. This is relevant, because executives could understand if higher scores in, for example CRM data task utility, lead to a template heuristic orientation. Likewise, executives should make critical assessments about the exhibited information search behavior patterns versus desired patterns. In particular, executives should monitor the intensity of scanning orientation in the marketing group. Probably, this group should be the lead organization exhibiting this behavior.

REFERENCES

Abelson, R. 1979 Differences between belief systems and knowledge systems. *Cognitive Science, 3*, 355-66

Abril, R.M. 2001, "An assessment of the theory underpinning the role of information quality in the single-loop decision making model". *Sixth International Conference on Information Quality*, Pierce, E.M. *et* al. eds., Cambridge, MA: 273-89

Abril, R.M. 2002, "An application of the Project Management Body of Knowledge to a theory-testing research". *International Symposium on Research Methods*, Loch, K. *et al.* eds., Las Vegas, NV

Abril, R.M. 2002, "Procedural rationality as a framework for assessing case research methods". *International Symposium on Research Methods*, Hackney, R. *et* al. eds., Barcelona, Spain

Abualsamh, R.A., Carlin, B. and McDaniel Jr., R.R. 1990 Problem structuring heuristics in strategic decision making. *Organizational behaviour and human decision processes, 45*, 159-74

Adams, D.A., Nelson, R.R. and Todd, P.A. 1992 Perceived uselfulness, ease of use and Usage of IT: A replication. *Management Information Systems Quarterly, 16* 2 , 227-47

Adriaenssens, C. and Cadman, L. 1999 An adaptation of moderated e-mail focus groups to assess the potential for a new online Internet financial services offer in the UK. *Journal of Market Research Society, 41* 4 , 417-24

Agarwal, C.C. and Parthasarathy, S. 2001, "Mining massively incomplete data sets by conceptual reconstruction". *Seventh Int. Conf. Knowledge Discovery and Data Mining*, anonymous eds., 227-32

Agarwal, R., Sinha, A.P. and Tanniru, M. 1996 Cognitive fit in requirements modeling: A study of object and process methodologies. *Journal of Management Information Systems, 13* 2 , 137-62

Agosta, L. 2002 Giga data warehousing survey validates expectations, contains surprises. Giga Information Group, Inc.

Agosta, L. 2002 Market Overview 2002: Information Quality. Giga Information Group, Inc.

Agosta, L. 2003 Data warehousing first half 2003: Data mining and data quality developments. Giga Research

Aguilar, F.J. 1967 *Scanning the business environment.* NY: Macmillan

Ahire, S.L., Golhar, D.Y. and Waller, M.A. 1996 Development and validation of TQM implementation constructs. *Decision Sciences, 27* 1

Akgun, A.E. 2001 Learning typology for new product development teams: A socio-cognitive perspective. Stevens Institute of Technology

Aldag, R.J. and Power 1986 An empirical assessment of computer-assisted decision analysis. *Decision Sciences, 17* 4 , 572-88

Anderson, E.W. and Sullivan, M.W. 1993 The antecedents and consequences of customer satisfaction for firms. *Marketing Science, 12*, 125-43

Anderson, J.C. and Gerbing, D.W. 1988 Structural equation modelling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103* 3, 411-23

Anderson, J.C. and Narus, J.A. 1990 A model of distributor firm and manufacturer firm working partnerships. *Journal of Marketing, 48* Fall, 62-74

Anderson, J.C. and Narus, J.A. 1996 Rethinking distribution: Adaptive channels. *Harvard Business Review, 74* July/August, 112-22

Anderson, J.R. 1983 *The architecture of cognition.* Cambridge, MA: Harvard University Press

Anderson, J.R. 1993 *Rules of the mind.* Hillsdale, NJ: Lawrence Erlbaum Associates

Arbuckle, J.L. 2003 Software AMOS 5.0

Argyris, C. 1976 Single-loop and double-loop models in research on decision-making. *Administrative Science Quarterly, 21* 3 , 363-75

Argyris, C. and Schön, D.A. 1978 *Organizational learning: A theory of action perspective.* Reading, MA: Addison Wesley

ASA 1999 Ethical guidelines for statistical practice. American Statistical Association

Asheghian, P. and Ebrahimi, P. (1990) International business: Economics, environment, and strategies. New York: Harper Collins

Averett, R.L. 1991 The role of cognitive style in the use of information for organizational decision-making. GEORGE MASON UNIVERSITY

Bagozzi, R.P. 1996 *Measurement in marketing research: Basic principles of questionnaire design.* Cambridge, MA: Blackwell Publishers

Bagozzi, R.P., Yi, Y. and Phillips, L. W. 1991 Assessing construct validity in organizational research. Administrative Science Quarterly, 36 3, 421-58

Bagranoff, N.A., Houghthon, K.A. and Hronsky, J. (1994) The structure of meaning in accounting: A cross-cultural experiment. Behavioral Research in Accounting, 6 (Supplement) , 35-57

Bailey, A. and Johnson, G. 1995, "The processes of strategy development" in *The CIMA Handbook of Strategic Management*, Thompson, J.L. eds., Oxford: Butterworth Heinemann, 71-98

Bailey, J.E. and Pearson, S.W. 1983 Development of a tool for measuring and analyzing computer user satisfaction. *Management Science, 29* 5 , 530-45

Bajaj, A. 2000 SMMM: A Metric Based Framework to Evaluate the Scalability of Multiple Model Methodologies. Carnegie Mellon. TheHeinz School

Bandura, A. 1977 Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215

Bandura, A. 1982 Self-efficacy mechanism in human agency. *American Psychologist, 37*, 122-47

Bandura, A. 1986 *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall

Bandura, A. 1994, "Self-efficacy" in *Encyclopedia of human behavior*, vol. 4, Ramachaudran, V.S. eds., NY: Academic Press, 71-81

Bandura, A. 1997 *Self-efficacy: The exercise of control.* NY: Freeman

Bandura, A. 2001 Guide for constructing self-efficacy scales. Stanford University

Barki, H. and Huff, S.L. 1985 Change, attitude to change, and decision support system success: An empirical assessment. *Information & Management, 9*, 261-8

Baron, R.M. and Kenny, D.A. 1986 The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-82

Beach, L.R. and Mitchell, T.R. 1978 A contingency model for the selection of decision strategies. *The Academy of Management Review, 3* 3, 439-49

Beal, R.M. 2000 Competing effectively. Environmental scanning competitive strategy, and organizational performance in small manufacturing firms. *Journal of Small Business Management, 38* 1, 27-47

Belkin, N.J. 1980 Anomalous states of knowledge as a basis for information retrieval. *The Canadian Journal of Information Science, 5*, 133-43

Belkin, N.J., Oddy, R.N. and Brooks, H.M. 1982 ASK for information retrieval: part I. Background and theory. *Journal of Documentation, 38* June, 61-71

Belkin, N.J., Oddy, R.N. and Brooks, H.M. 1982 ASK for information retrieval: part II. Results of a design study. *Journal of Documentation, 38* September, 145-64.

Bell, D.A. 1993 From data properties to evidence. *IEEE Transactions on Knowledge and Data Engineering, 5* 6, 965-9

Benbasat, I., Dexter, A.S. and Masulis, P.S. 1981 An experimental study of the human/computer interface. *Communications of the ACM, 24* November, 752-62

Benbasat, I. and Schroeder, R.G. 1977 An experimental investigation of some MIS design variables. *Management Information Systems Quarterly, 1*, 37-50

Benbasat, I. and Taylor, R.N. 1982 Behavioral aspects of information processing for the dessign of management information systems. *IEEE Transactions on Systems, Man and Cybernetics, 12* 4, 439-50

Berger, C.R. 1979, "Beyond initial understanding:Uncertainty, understanding, and the development of interpersonal relationships" in *Language and social psychology*, Giles, H. and St. Clair R. N. eds., Oxford: Basil Blackwell, 122-44

Berscheid, E. 1994 Interpersonal relationships. *Anual Review of Psychology, 45*, 79-129

Blumstein, P. and Kollock, P. 1988 Personal relationships. *Anual Review of Sociology, 14*, 467-90

Bontis, N. and Fitz-enz, J. 2002 Intellectual capital ROI: A causal map of human capital antecedents and consequents. *Journal of Intellectual Capital, 3* 3, 223-47

Bourgeois III, L.J. and Eisenhardt, K.M. 1988 Strategic decision processes in high velocity environments: four cases in the microcomputer industry. *Management Science, 34* 7 , 816-35

Boynton, A.C. 1987 Predictors of managerial scanning. THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

Boynton, A.C. and Zmud, R.W. 1987 Information technology planning in the 1990's: Directions for practice and research. *Management Information Systems Quarterly, 11* 1, 59-71

Brannick, T. 1998 The marketing task-information interface: The Daft-MacIntosh information processing model revisited. *Irish Marketing Review, 11* 1, 27-38

Brannick, T. 2000 The enigma of marketing information. *IBAR- Journal of the Irish Academy of Management, 21* 2, 81-106

Bratko, I. 1994, "Learning and Noise". *MLnet Summer School on Machine Learning and Knowledge Acquisition*, anonymous eds., Dourdan, France: 121-43

Brazdil, P. and Clark, P. 1990, "Learning from imperfect data" in *Machine learning, meta-reasoning, and logics*, Brazdil, P. and Konolige, K. eds., Boston, MA: Kluwer Academic Publishers, 207-32

Brewer, M. and Richards, B. 2001 Segmentation by Behaviour - New Approaches to Managing Customer Relationships. *International Journal of Customer Relationship Management*, June/Julyl

Brewer, W.F. and Treyens, J.C. 1981 Role of schemata in memory for places. *Cognitive Psychology, 13*, 207-30

Bruner, J. 1990 *Acts of meaning*. Cambridge, MA: HARVARD UNIVERSITY.

Burke, K. 1969 *A grammar of motives*. Berkeley, CA: University of California Press

Burke, K. 1989 *On symbols and society*. Chicago, IL: University of Chicago Presss

Burky, L.B. 1990 Information evaluation: Individual process determinants in a DSS environment. UNIVERSITY OF PITTSBURGH

Butler, K.A. 1996 Usability engineering turns 10. *Interactions*, 3, 59-75

Butler, R.J. 1996, "Decision making" in *International Encyclopedia of Business and ManagementInternational Encyclopedia of Business and Management*, Warner, M. eds, 951-1001

Byrne, B.M. 2001 *Structural equation modeling with AMOS. Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates

Campbell, J.P. 1994, "Modeling the performance prediction problem in industrial and organizational psychology" in *Handbook of industrial and organizational*

psychology, vol. 1, Dunnette, M.D. and Hough L. M. eds., Palo Alto, CA: Consulting Psychologists Press

Chapman, R.G. 1989 Problem-Definition in marketing research studies. *The Journal of Services Marketing*, *3* 3, 51-9

Chen, C., Lee, S.-Y. and Stevenson, H.W. (1995) Response style and cross-cultural comparisons of rating scales among Est Asian and North American students. Psychological Science, 6 (3), 170-5

Chen, L., Soliman, K.S., Mao, E. and Frolick, M.N. 2000 Measuring user satisfaction with data warehouses: an exploratory study. *Information & Management*, *37* 3, 103-10

Chi, T. and Fan, D. 1997 Cognitive limitations and investment "myopia". *Decision Sciences*, *28* 1, 27-57

Chiang, R.H.L., Barron, T.M. and Storey, V.C. 1994 Reverse engineering of relational databases: Extraction of an EER model from a relational database. *Data & Knowledge Engineering*, *12*

Chin, W., W., Marcolin, B., L. and Newsted, P., R. 1996, "A partial least squares latent variable modeling approach for measuring ineractin effects: Results from Monte Carlo simulation study and voice mail emotion/adoption study". *Seventeenth International conference on information systems*, DeGross, J.I. *et* al. eds., Cleveland, OH

Choo, C.W. 1999 The art of scanning the environment. *Bulletin of the American Society for Information Science*, *25* 3, 21-4

Choo, C.W. 1993 Environmental scanning: Acquisition and use of information by chief executive officiers in the Canadian telecommunictions industry. University of Toronto

Chorba, R.W. and New, J.L. 1980 Information support for decision-maker learning in a competitive environment: An experimental study. *Decision Sciences*, *11*, 603-15

Chung, J. 1995 Auditor's confidence and the audit expectation gap. *Australian Accountant*, June , 26-30

Churchill Jr., G.A. 1979 A paradigm for developping better measures of marketing constructs. *Journal of Marketing Research, 16*, 64-73

Churchill Jr., G.A. 1991 *Marketing research. Methodological foundations*. The Dryden Press

Churchman, C.W. 1971 *The Design of inquiring systems: Basic concepts of systems and organization*. NY: Basic Books, Inc.

Codd, E.F. 1970 A relational model of data for large shared databases. *Communications of the ACM, 13* 6, 377-87

Codd, E.F. 1972, "Further normalization of the data base relational model" in *Data base systems*, vol. 6, Rustin, R. eds., Englewood Cliffs, NJ: Prentice-Hall, 33-64

Cohen, M.D., March, J.G. and Olsen, J.P. 1972 A garbage can model of organizational choice. *Administrative Science Quarterly, 17*, 1-25

Cohen, M.S. 1995, "The naturalistic basis of decision biases" in *Decision making in action: Models and methods*, Klein, G.A., Orasanu J., Calderwood R., and Zsambok C. E. eds., Norwood, NJ: Ablex, 51-99

Compeau, D.R., Higgins, C.A. and Huff, S.L. 1999 Social cognitive theory and individual reactions to computing technology:A longitudinal study. *Management Information Systems Quarterly, 23* 2, 145-58

Corey, R.J. and Wilson, D.T. 1994 Using motivation as a basis for understanding and improving direct marketing relationships. *Journal of Direct Marketing, 8* 4, 28-39

Cowan, D.A. 1986 Developing a process model of problem recognition. *The Academy of Management Review, 11* 4, 763-76

Cowan, D.A. 1988 Executive's knowledge of organizational problem types: Applying a contingency perspective. *Journal of Management, 14*, 513-27

Cravens, D.W. 1998 Implementation strategies in the market-driven strategy era. *Journal of the Academy of Marketing Science, 26* 3, 237-41

Cronbach, L.J. and Meehl, P.E. 1955 Construct validity in psychological tests. *Psychological Bulletin, 52* 4, 281-302

Curasi, C.F. 2001 A critical exploration of face-to-face interviewing vs. computer-mediated interviewing. *International Journal of Marketing Research, 43* 4, 361-75

Curwin, J. and Slater, R. 1996 *Quantitative Methods for Business Decisions.* International Thomson Business Press.

Cyert, T. and March, J.G. 1963 *The behavioral theory of the firm.* Englewood Cliffs, NJ: Prentice-Hall

Daft, R.L. and Lengel, R.H. 1984 Information richness: A new approach to managerial behavior and organization design. *Research in Organizational Behavior, 6*, 191-233

Daft, R.L. and Lengel, R.H. 1986 Organizational information requirements, media richness and structural design. *Management Science, 32* 5, 554-71

Daft, R.L. and Macintosh, N.B. 1981 A tentative exploration into the amount and equivocality of information processing in organizational work units. *Administrative Science Quarterly, 26* 2, 207-24

Daft, R.L. and Weick, K.E. 1984 Toward a model of organizations as interpretation systems. *The Academy of Management Review, 9* 2, 284-95

Davenport, T.H., Eccles, R.G. and Prusak, L. 1992 Information politics. *Sloan Management Review, 34* 1, 53-65

Davenport, T.H., Harris, J.G. and Kohli, A.K. 2001 How do they know their customers so well? *Sloan Management Review, 42* 2, 63-73

Davidson, J.E., Deuser, R. and Sternberg, R.J. 1994, "The role of metacognition in problem solving" in *Metacognition : Knowing about knowing*, Metcalfe, J. and Shimamura, A.P. eds., Cambridge, MA: MIT press, 207-26

Davis, F.D. 1989 Perceived usefulness, perceived ease of use and user acceptance of IT. *Management Information Systems Quarterly, 13* 3, 318-41

Davis, F.D., Bagozzi, R.P. and Warshaw, P.R. 1989 User acceptance of computer technology: A comparison of two theoretical models. *Management Science, 35* 8, 982-1003

Day, G.S. 1994 The capabilities of market-driven organizations. *Journal of Marketing, 58* October , 37-52

Day, G.S. and Nedungadi, P. 1994 Managerial representations of competitive advantage. *Journal of Marketing, 58* April, 31-44

De Geus, A.P. 1988 Planning as learning. *Harvard Business Review, 66* 2, 70-4

De Wulf, K., Odekerken-Schroder, G. and Lacobucci, D. 2001 Investments in consumer relationships: A cross-country and cross-industry exploration. *Journal of Marketing, 65* 4, 33-50

Dean Jr., J.W. and Sharfman, M.P. 1993 Procedural rationality in the strategic decision-making process. *Journal of Management Studies, 30* 4, 587-610

DeLone, W.H. and McLean, E.R. 1992 Information systems success: The quest for the dependent variable. *Information Systems Research, 3* 1, 60-95

DeLorme, D.E., Zinkhan, G., M. and French, W. 2001 Ethics and the internet: Issues associated with qualitative research. *Journal of Business Ethics, 33*, 271-86

Dervin, B. 1992, "From the mind's eye of the user: The sense-making qualitative-quantitative methodology" in *Qualitative research in information management*, Glazier, J.D. and Powell R. R. eds.

Dervin, B. and Nilan, M. 1986, "Information needs and uses." in *Annual review of information science and Technology*, Williams, M. eds., Medford, NJ: Learned Information, 3-33

Dery, D. 1983 Decision-making, problem-solving, and organizational learning. *OMEGA, 11*, 321-8

Deshpande, R., Farley, J.U. and Webster Jr., F.E. 1993 Corporate culture, customer orientation, and innovativeness in Japanese firms: A quadrad analysis. *Journal of Marketing, 57* January, 23-37

Deshpande, R. and Zaltman, G.R. 1982 Factors affecting the use of market research information: A path analysis. *Journal of Marketing Research, 19* 1, 14-31

Deshpande, R. and Zaltman, G.R. 1984 A comparison of factors affecting researcher and manager perceptions of market research use. *Journal of Marketing Research, 21* February, 32-8

Dillman, D. and Bowker, D.K. 2001, "The web questionnaire challenge to survey methodologists" in *Dimensions of internet science*, Reips, U.-D. and Bosnjak M. eds., Lengerich, Germany: Pabst Science Publishers

Dillman, D.A. and Carley-Baxter, L. 2000, "Structural determinants of mail survey response rates over a 12 year period: 1988~1999". *2000 proceedings of the survey research methods section*, anonymous eds., Alexandria, VA: American Statistical Association 394-9

DMA 2003 The DMA 2003 response rate study. The Direct Marketing Association

Doig, G. 2002 A study of the factors that influence the perceptions of strategic information value in major UK based organisations. Loughborough University

Dutton, J.E., Fahey, L. and Narayanan, U.K. 1983 Toward understanding strategic issue diagnosis. *Strategic Management Journal, 4*, 307-23

Dutton, J.E. and Jackson, S.E. 1987 Categorizing strategic issues: Links to organizational actions. *Academy of Management Journal Review, 12*, 76-90

Dweck, C.S. and Elliot, E.S. 1983, "Achievement motivation" in *Handbook of child psychology: Socialization, personality and social development*, vol. 4, Mussen, P.H. and Heatherington E. M. eds., 644-91

Easterby-Smith, M., Thorpe, R. and Lowe, A. 1991 *Management research. An introduction.* SAGE

Eckerson, W. 2001 Data quality and the bottom line. Achieving business success through a commitment to high quality data. The Data Warehousing Institute

Emery, F.E. and Trist, E.L. 1965 The causal texture of organizational environments. *Human Relations, 18*, 21-32

English, L.P. 1999 *Improving data warehouse and business information quality: Methods for reducing costs and increasing profits.* John Wiley & Sons

Eppler, M.J. and Wittig, D. 2000, "Conceptualizing Information Quality: A Review of Information Quality Frameworks from the Last Ten Years". *TDQM*, anonymous eds., Cambridge, MA

Evans, J.S.B.T. 1989 *Biases in human reasoning: Causes and consequences.* Hove, UK: Lawrence Erlbaum Associates.

Fahy, J. 1998 Improving response rates in cross-cultural mail surveys. *Industrial Marketing Management, 27* 6, 459-66

Fedorowicz, J. and Lee, Y.W. 1998, "Accounting Information Quality: Reconciling Hierarchical and Dimensional Contexts". *TDQM*, anonymous eds., Cambridge, MA

Feist, G. 1994 Personality and working style predictors of integrative complexity: A study of scientists' thinking about research and teaching. *Journal of Personality and Social Psychology, 67*, 474-84

Feldman, M.S. and March, J.G. 1981 Information in organization as signal and symbol. *Administrative Science Quarterly, 26* 2, 171-86

Fiol, C.M. and Lyles, M.A. 1985 Organizational learning. *Academy of Management Review, 10* 4, 803-13

Fishbein, M. 1967, "Attitude and the prediction of behaviour" in *Readings in attitude theory and measurement*, Fishbein, M. eds., NY: Wiley, 477-92

Fishbein, M. and Ajzen, I. 1975 *Belief, attitude, intention and behaviour: An introduction to theory and research.* Reading, MA: Addison Wesley

FJC 2000 Reference manual on scientific evidence. Federal Judicial Center

Fletcher, G.J.O., Danilovics, P., Fernandez, G., Peterson, D. and Reeder, G.D. 1986 Attributional complexity: An individual difference measure. *Journal of Personality and Social Psychology, 54* 4, 875-84

Fletcher, K. 1995 *Marketing management and information technology.* Prentice-Hall

Folkes, V. 1988 The availability heuristic and perceived risk. *Journal of Consumer Research, 15*, 3-23.

Folkman, S. and R.S. Lazarus 1988 Manual for the Ways of Coping Questionnaire. Consulting Psychologists Press

Fornell, C. and Larcker, D.F. 1981 Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18* 1, 39-50

Foss, B., Henderson, I., Johnson, P., Murray, D. and Stone, M. 2002 Managing the quality and completeness of customer data. *Journal of Database Marketing, 10* 2, 139-58

Fox, C., Levitin, A. and Redman, T.C. 1994 The notion of data and its quality dimensions. *Information Processing & Management, 30* 1

Franz, H. 1999 How computer mediated communication affects information overload in distributed systems. Brunel University

Fredrickson, J.W. 1984 The comprehensiveness of strategic decision processes: Extension, observations, future directions. *Academy of Management Journal, 27* 3, 445-66

Gallagher, C.A. 1974 Perceptions of the value of a management information system. *Academy of Management Journal, 17* 1, 46-55

Gatian, A.W. 1994 Is user satisfaction a valid measure of system effectiveness? *Information & Management, 26* 3, 119-31

Gile, K. 2002 Developing a business intelligence strategic plan. Giga Information Group, Inc.

Gist, M.E. and Mitchell, T.R. 1992 Self-efficacy: A theoretical analysis of its determinants and malleability. *The Academy of Management Review, 17* 2, 183-211

Glazer, R. 1991 Marketing in an information-intensive environment: Strategic implications of knowledge as an asset. *Journal of Marketing, 55* 4, 1-18

Glazer, R. 1993 Measuring the value of information: The information-intensive organization. *IBM Systems Journal, 32* 1, 99-110

Goia, D.A. 1986, "Symbols, scripts, and sense-making: Creating meaning in the organizational experience" in *The thinking organization: Dynamics of organizational social cognition*, Sims, H.P. and Gioia D. A. eds., San Francisco: Jossey-Bass, 49-74

Gonzalez, G.R. 2001 The role of information acquisition and knowledge use in managing customer relationships. ARIZONA STATE UNIVERSITY

Goodhue, D.L., Klein, B.D. and March, S.T. 2000 User evaluations of IS as surrogates for objective performance. *Information & Management, 38* 2, 87-101

Goodhue, D.L., Wybo, M.D. and Kirsch, L.J. 1992 The impact of data integration on the cost and benefits of IS. *Management Information Systems Quarterly, 16* 3, 293-311

Goodhue, D.L. and Thompson, R.L. 1995 Task-technology fit and individual performance. *Management Information Systems Quarterly, 19* 2, 213-36

Goslar, M.D., Green, G.I. and Hughes, T.H. 1986 Decision support systems:An empirical assessment for decision making. *Decision Sciences, 17* 1, 79-91

Gottschalk, P. 1998 Content characteristics of formal information technology strategy as implementation predictors. Brunel University

Goul, M., Shane, B. and Tonge, F.M. 1986 Using a knowledge-based decision support system in strategic planning decisions: An empirical study. *Journal of Management Information Systems, 2* 4, 70-84

Graham, C., N. Latimer, F. Biscotti, J. Correia, C. Eschinger, C. Pang and T. Topolinski 2004 Software market research methodology and definitions, 2003-2004. Gartner

Griffin, A. and Hauser, J.R. 1996 Integrating R&D and marketing: A review and analysis of the literature. *The Journal of Product Innovation Management, 13* 3, 191-215

Grönroos, C. 1995 Relationship marketing:The strategy continuum. *Journal of the Academy of Marketing Science, 23* 4, 252-4

Grooms, T.F. 2001 Marketing intelligence: Executive management perceptions of value. Henley Management College.

Gwinner, K.P., Gremler, D.D. and Bitner, M.J. 1998 Relational benefits in services industries: The customer's perspective. *Journal of the Academy of Marketing Science, 26* 2, 101-14

Hainaut, J.-L. , Henrard, J., Roland, D., Englebert, V. and Hick, J.-M. 1996, "Structure elicitation in database reverse engineering". *3rd Working Conference on Reverse Engineering WCRE '96*, anonymous eds., Monterey, CA

Hair Jr., J.F. , Anderson, R.E. , Tatham, R.L. and Black, W.C. 1998 *Multivariate data analysis*. Prentice-Hall

Hair Jr., J.F. , Babin, B. , Money, A.H. and Samouel, P. 2003 *Essentials of business research methods*. Wiley

Haley, B.J. 1997 Implementing the decision support infraestructure: Key success factors in data warehousing. UNIVERSITY OF GEORGIA

Harris, M. and Schaubrock, J. 1990 Confirmatory modeling in organization behavior/human resource management: Issues and applications. *Journal of Management, 16*, 337-60

Hendrickson, A.R., Massey, P.D. and Cronan, T.P. 1993 On the test-retest reliability of perceived usefulness and perceived ease of use scales. *Management Information Systems Quarterly, 17* 2, 227-34

Henson, R.K., Kogan, L.R. and Vacha-Haase, T. 2001 A reliability generalization study of the teacher efficacy scale and related instruments. *Educational and Psychological Measurement, 61* 3, 404-20

Heppner, P.P. 1988 *The problem solving inventory manual.* Palo Alto, CA: Consulting Psychologist Press

Heppner, P.P., Cook, S.W., Wright, D.M. and Johnson Jr., W.C. 1995 Progress in resolving problems: A problem-focussed style of coping. *Journal of Counseling Psychology, 42* 3 279

Heppner, P.P., Cooper, C., Mulholland, A. and Wei, M. 2001 A brief, multidimensional, problem-solving psychotherapy outcome measure. *Journal of Counseling Psychology, 48* 3, 330-43

Herschel, G. 2002 CRM analytics suite key criteria for 2002. Gartner Research

Hodgkinson, G.P. and Johnson, G. 1994 Exploring the mental models of competitive strategists: The case for a processual approach. *The Journal Of Management Studies, 31* 4, 525-51

Hogarth, R.M. 1987 *Judgement and choice: The psychology of decisions.* Chichester: John Wiley & Sons

Hogarth, R.M. and Makridakis, S. 1981 Forecasting and planning: An evaluation. *Management Science, 27* 2, 115-38

Homburg, C. and Pflesser, C. 2000 A multiple-layer model of market-oriented organizational culture: Measurement issues and performance outcomes. *Journal of Marketing Research, 37* November, 449-62

Hough and Paullin, C. 1994, "Construct-oriented scale construction: The rational approach" in *The biodata handbook. Theory, research, and use of biographical information in selection and performance prediction*, Stokes, G.S., Mumford M. D., and Owens W. eds., 109-45

Huang, K.-T. , Lee, Y.W. and Wang, R.Y. 1998 *Quality information and knowledge management*

Huber, G.P. 1991 Organizational learning: The contributing processes and the literatures. *Organization Science, 2* 1, 88-115

Huber, G.P. and Daft, R.L. 1987, "The information environments in organizations" in *Handbook of organizational communication*, Jablin, F. .M., Putnam L. L., Roberts K. H., and Porter L. W. eds., Newbury Park, CA: Sage, 130-64

Inmon, W.H. 1996 *Building the data warehouse.* NY: John Wiley & Sons

Jabri, M.M. 1991 The developmenet of conceptually independent subscales in the measurement of modes of problem solving. *Educational and Psychological Measurement, 51* 2, 975-83

Jackson, B.B. 1985 *Winning and keeping industrial customers: The dynamics of customer relationships.* Lexington, Mass.: Heath

Jarke, M., Jeusfeld, M.A., Quix, C. and Vassiliadis, P. 1999 Architecture and quality in data warehouses: An extended repository approach. *Information Systems, 24* 3, 229-53

Jaworski, B.J. and Kohli, A.K. 1993 Market orientation: Antecedents and consequences. *Journal of Marketing, 57* July, 53-70

Jenkins, M. and J. Ricketts 1985 The development of a MIS satisfaction questionnaire: An instrument for evaluating user satisfaction with turnkey decision support systems. Indiana University

Jobber, D. and O'Reilly, D. 1998 Industrial mail surveys. *Industrial Marketing Management, 27* 2, 95-107

Jobber, D. and Watts, M. 1988 User attitudes towards marketing information systems. A UK survey of manufacturing companies. *Marketing Intelligence & Planning, 6* 2, 30-5

Jones, J.W. and McLeod Jr., R. 1986 The structure of executive information systems: An exploratory analysis. *Decision Sciences, 17* 2, 220-49

Johns, S.K., Smith, L.M. and Strand, C.A. (2002) How culture affetcs the use of information technology. Accounting Forum, 27 (1) , 84-109

Kahneman, D. and Tversky, A. 1972 Subjective probabiliy: A judgement of representativeness. *Cognitive Psychology, 3* 3, 430-54

Kanfer, R. and Ackerman, P.L. 1989 Motivation and cognitive abilities: An integrative/aptitude treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*, 657-89

Kaplan, C. and Simon, H.A. 1990 In search of insight. *Cognitive Psychology, 22*, 374-419

Keillor, B., Owens, D. and Pettijohn, C. (2001) A cross-cultural/cross-national study of influencing factors and socially desirable response biases. International Journal of Market Research, 43 (1), 63-84

Keller, K.L. 1993 Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing, 57*, 1-22

Kelley, H.H. 1967, "Attribution theory in social psychology". *Nebraska symposium on motivation*, Levine, D. eds., Lincoln, NE: University of Nebraska Press 192-241

Kelley, H.H., Berscheid, E., Christensen, A. , Harvey, H.H. , Huston, T.L. , Levinger, G. , McClintoc, E. , Peplau, L.A. and Peterson, D. 1983 *Close relationships*. San Francisco: Freeman

Kelly, S. 1997 *Data Warehousing in action*. Chichester: John Wiley & Sons

Kettinger, W.J. and Lee, C.C. 1994 Perceived service quality and user satisfaction with the information services function. *Decision Sciences, 25* 5/6

Kiesler, S. and Sproull, L. 1982 Managerial response to changig environments: Perspectives on problem sensing from social cognition. *Administrative Science Quarterly, 27*, 548-70

Kim, K.K. 1989 User satisfaction: A synthesis of three different perspectives. *Journal of Information Systems, 6* 4, 1-12

King, W.R. and Epstein, B.J. 1983 Assessing information system value: An experimental study. *Decision Sciences, 14* 1

Kirkman, B.L. and Rosen, B. 1999 Beyond self-management: Antecedents and consequences of team empowerment. *Academy of Management Journal, 42* 1, 58-64

Klein, H.J. 1989 An integrated control theory model of work motivation. *Academy of Management Review, 14* 2, 150-72

Kleysen, R.F. and Street, C. 2001 Toward a multi-dimenesional measure of individual innovative behavior. *Journal of Intellectual Capital, 2* 3, 284-96

Knuth, D.E. 2001 *Things a computer scientist rarely talks about*. Center for the Study of Language and Information

Kolb, D.A. 1974, "On management and the learning process" in *Organizational Psychology*, Kolb, D.A., Rubin I. M., and MacIntyre J. M. eds., 27-42

Kolb, D.A. 1984 *Experiential Learning. Experience as the Source of Learning and Development*. Prentice-Hall

Kotler, P. 1988 *Marketing management. Analysis, planning, implementation, and control.* Englewood Cliffs, NJ: Prentice Hall International

Kuhlthau, C.C. 1993 *Seeking meaning: A process approach to library and information services.* Norwood, NJ: Ablex

Kuhlthau, C.C. 1999 Accommodating the user's information search process: Challenges for information retrieval system designers. *Bulletin of the American Society for Information Science, 25* 3, 12-6

Laffont, J.-J. 1989 *The economics of uncertainty and information.* Cambridge MA: The MIT Press

Lamberts, K. and Shanks, D.R. 1997 *Knowledge, concepts, and categories studies in cognition.* MIT

Lau, C. and Woodman, R.W. 1995 Understanding organizational change: A schematic perspective. *Academy of Management Journal, 38* 2, 537-54

Lau, R.R. and Redlawsk, D.P. 2001 Advantages and disanvantages of cognitive heuristics in political decision making Forthcoming. *American Journal of Political Science, Forthcoming*

Lawrence, P.R. and Lorsch, J.W. 1967 *Organization and environment.* Cambridge, MA: Harvard University Press

Lee, S., Courtney, J.F. and O'Keefe, R.M. 1992 A system for organizational learning using cognitive maps. *OMEGA, 20* 1, 23-36

Leidner, D.L., Carlsson, S., Elam, J. and Corrales, M. 1999 Mexican and Swedish managers' perceptions of the impact of EIS on organizational intelligence, decision making and structure. *Decision Sciences, 30* 3

Leigh, T.W. and Marshall, G.W. 2001 Research priorities in sales strategy and performance. *The Journal of Personal Selling & Sales Management, 21* 2, 83-93

Levitt, B. and March, J.G. 1988 Organizational learning. *Anual Review of Sociology, 14* February, 319-40

Liang, T.-P. 1986 Critical success factors of decision support systems: An experimental study. *Data Base*, Winter, 3-16

Lipshitz, R. and Strauss, O. 1997 Coping with uncertainty: A naturalistic decision-making analysis. *Organizational Behavior and Human Decision Processes, 69* 2, 149-63

Little, R.J.A. and Rubin, D.B. 2002 *Statistical analysis with missing data*. Wiley

Low, G.S. and Mohr, J.J. 2001 Factors affecting the use of information in the evaluation of marketing communications productivity. *Journal of the Academy of Marketing Science, 29* 1, 70-88

Lucas Jr., H.C. 1973 User reactions and the management of information services. *Management Informatics, 2* 4, 165-72

Lucas Jr., H.C. 1981 An experimental investigation of the use of computer-based graphics in decision making. *Management Science, 27* 7, 757-68

Lucas Jr., H.C. and Nielson, N.R. 1980 The impact of the mode of information presentation on learning and performance. *Management Science, 26* 10, 982-93

Lyles, M.A. and Mitroff, I.I. 1980 Organizational problem formulation: An empirical study. *Administrative Science Quarterly, 25* 1, 102-19

Mahmood, M.A. and Medewitz, N. 1985 Impact of design methods on decision support systems success: An empirical assessment. *Information & Management, 9* 3, 137-51

Maier, J.L., Rainer Jr., R.K. and Snyder, C.A. 1997 Environmental scanning for information technology: An empirical investigation. *Journal of Management Information Systems, 14* 2, 177-200

Maltz, E. and Kohli, A.K. 1996 Market intelligence dissemination across functional boundaries. *Journal of Marketing Research, 33* February, 47-61

Mann, L., Radford, M., Burnett, P., Ford, S., Bond, M., Leung, K., Nakamura, H., Vaughan, G. and Yang (1998) Cross-cultural differences The Melbourne decision making questionnaire: An instrument for measuring patterns for coping with decisional conflictin self-reported decision-making style and confidence. International Journal of Psychology, 33 (5), 325-35.

Maoz, M., R. DeSisto, C. Marcus, G. Herschel, E. Kolsky, E. Thompson and T. Berg 2003 The Gartner CRM software vendor guide: 2003. Gartner Research

Marakas, G.M. 1998 *Decision support systems in the 21st century*. Prentice-Hall

March, J.G. and Olsen, J.P. 1976, "Organizational choic under ambiguity" in *Ambiguity and choice in organizations*, March, J.G. and Olsen J. P. eds., Bergen: Universitestforlaget, 10-23

March, J.G. and Simon, H.A. 1958 *Organizations*. NY: Wiley.

Marshall, C. and Rossman, G. 1995 *Designing qualitative research*. Thousand Oaks, CA: SAGE.

Massey, A.P. and O'Keefe, R.M. 1993 Insights from attempts to validate a multi-attribute model of problem definition quality. *Decision Sciences, 24* 1

McDonough, A.M. 1963 *Information economics and management systems*. NY: McGraw-Hill

McFadzean, E. 1996 New ways of thinking: An evaluation of thesis K-groupware and creative problem solving. Brunel University

McGrath, J.E. 1979, "Toward a 'Theoy of Method' for research on organizations" in *Reseach in organizations: Issues and controversies*, Mowday, R.T. and Steeres R. M. eds., Santa Monica, CA: Goodyear Publishing, 4-21

McGrath, J.E. 1982, "Dilematics. The study of research choices and dilemmas" in *Judgement calls in research*, McGrath, J.E., Martin J., and Kulka R. A. eds. SAGE

McKnight, W. 2000 The CRM-ready data warehouse. *DM Review*, Februery

Menon, A. and Varadarajan, P.R. 1992 A model of marketing knowledge use within firms. *Journal of Marketing*, *56* 4, 53-71

Miles, M.B. and Huberman, A.M. 1994 *Qualitative data analysis*. Sage Publications

Miles, R.E. and Snow, C.C. 1978 *Organizational strategy, structure and process*. NY: McGraw-Hill

Miller, D.T. and Turnbull, W. 1986 Expectancies and interpersonal processes. *Anual Review of Psychology*, *37*, 233-56

Miller, G.A. 1956 The magical number seven, plus or minus two: Some limits to our capability for processing information. *Psychological Review*, *63* 2, 81-97

Miller, G.A., Galanter, E. and Pribram, K.H. 1960 *Plans and the structure of behaviour*. NY: Holt, Rinehart and Winaston, Inc.

Mintzberg, H. 1983 *Structure in fives: Designing effective organizations*. Englewood Cliffs, NJ: Prentice-Hall

Mintzberg, H., Raisinghani, D. and Théorêt, A. 1976 The structure of unstructured decision processes. *Administrative Science Quarterly*, *21* 2, 246-75

Mitroff, I. and Featheringham, I. 1974 On systematic problem solving and the error of the third kind. *Behavioral Science*, *19* 6, 383-93

Moenaert, R.K. and Souder, W.E. 1996 Context and antecedents of information utility at the R&D/Marketing interface. *Management Science*, *42* 11, 1592-610

Moorman, C. 1995 Organizational market information processes: Culture antecedents and new product outcomes. *Journal of Marketing Research*, *32* 3, 318-35

Moorman, C., Zaltman, G.R. and Deshpande, R. 1992 Relationships bewteen providers and users of market research. *Journal of Marketing Research, 29* 3, 314-28

Moorman, C. and J.R. Austin 1995 The paradox of low quality and high use: How researcher trust impacts market research outcomes. Marketing Science Institute

Morahan-Martin, J.M. (2004) How internet users find, evaluate, and use on-line health information: A cross-cultural review. CyberPsychology & Behavior, 7 (5), 497-510

Morgan, R.M. and Hunt, S.D. 1994 The commitment-trust theory of relationship marketing. *Journal of Marketing, 58* 3, 20-38

Muller, R. 2002 Communication of information technology project sponsors and managers in buyer-seller relationship. Brunel Unversity

Munro, M.C., Huff, S.L., Marcolin, B.L. and Compeau, D.R. 1997 Understanding and measuring user competence. *Information & Management, 33* 1, 45-57

Munro, M.C. and Davis, G.B. 1977 Determining management information needs: A comparison of methods. *Management Information Systems Quarterly, 1* 2, 55-67

Myers, M.D. 1997 Qualitative research in IS. *Management Information Systems Quarterly, 21* 2, 241-2

Narver, J.C. and Slater, S.F. 1990 The effect of market orientation on business profitability. *Journal of Marketing, 54* 4, 20-35

Neal, G. and Heppner, P.P. 1986 Problem-solving self-appraisal, awareness, and utilization of campus helping resources. *Journal of Counseling Psychology, 33* 1 39

Nevett, T. 1991 Historical investigation and the practice of marketing. *Journal of Marketing, 55* July, 13-23

Newell, A. and Simon, H.A. 1972 *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall

Nonaka, I. 1994 A dynamic theory of organizational knowledge creation. *Organization Science, 5* 1, 14-37

Orasanu, J. and Connolly, T. 1995, "The reinvention of decision making" in *Decision making in action: Models and methods*, Klein, G.A., Orasanu J., Calderwood R., and Zsambok C. E. eds., Norwood, NJ: Ablex, 3-20

O'Reilly III, C.A. 1982 Variations in decision makers' use of information sources: The impact of quality and accesibility of information. *Academy of Management Journal, 25* 4, 756-71

Orton, J.D. and Weick, K.E. 1990 Loosely coupled systems: A reconceptualization. *The Academy of Management Review, 15* 2, 203-23

Ostroff, C. and Kozlowski, S.W.J. 1992 Organizational socialization as a learning process: The role of information acquisition. *Personnel Psychology, 45*, 849-74

Oxford 1993 *The new shorter Oxford dictionary*. Clarendon Press

Pajares, F. 2002 Is self-efficacy represented in self-schemata? Emory University

Pajares, F. 2002 Self-efficacy beliefs and current directions in self-efficacy research. Emory University

Pajares, F., Hartley, J. and Valiante, G. 2001 Response format in writting self-efficacy assessment: Greater discrimination increases prediction. *Measurement and Evaluation in Counseling and Development, 33*, 214-21

Parker, M.M. , Benson, R.J. and Trainor, H.E. 1988 *Information economics. Linking business performance to information technology*. Prentice Hall International

Parker, S.K. and Axtell, C.M. 2001 Seeing another viewpoint: Antecedents and outcomes of employee perspective taking. *Academy of Management Journal, 44* 6, 1085-100

Pawlak, Z., Grzymala-Busse, J., Slowinski, R. and Ziarko, W. 1995 Rough sets. *Communications of the ACM, 38* 11 , 89-95

Payne, A. and Holt, S. 2001 Diagnosing customer value: Integrating the value process and relationship marketing. *British Journal of Management, 12*, 159-82

Payne, J.W. 1976 Task complexity and contingent processing in decision-making. An information search and protocol analysis. *Organizational Behavior and Human Performance, 16* 2, 366-87

Payne, J.W., Bettman, J.R. and Johnson, E.J. 1993 *The adaptive decision maker.* Cambridge, UK: Cambridge University Press

Peltier, J.W., Schibrowsky, J.A. and Davis, J. 1998 Using attitudinal and descriptive database information to understand interactive buyer-seller relationships. *Journal of Interactive Marketing, 12* 3, 32-45

Peynot, R. and Kinikin, E. 2003 Market update 2003: Packaged solutions for customer data integration. Giga Information Group, Inc.

Phillips, F. 2001 A research note on accounting students' epistemological beliefs, study strategies, and unstructured problem-solving performance. *Issues in Accounting Education, 16* 1, 21-39

Piatetsky-Shapiro, G. and Frawley, W.J. 1991 *Knowledge discovery in database.* AAAI Press

Pinsonneault, A. and Kraemer, K.L. 1993 Survey research methodology in management information systems: An assessment. *Journal of Management Information Systems, 10* 2, 75-105

Pitz, G.F., Heerboth, J. and Sachs, N.J. 1980 Procedures for eliciting choices in analysis of individual decisions. *Organizational behaviour and human performance, 26*, 396-408

Porter, M.E. 1980 *Competitive strategy: Techniques for analyzing industries and competitors.* NY: Free Press

Porter, S.S. and Inks, L.W. 2000 Cognitive complexity and salesperson adaptability: An exploratory investigation. *The Journal of Personal Selling & Sales Management, 20* 1, 15-21.

Pratch, L. and Jacobowitz, J. 1998 Integrative capacity and the evaluation of leadership: A multimethod assessment approach. *The Journal of Applied Behavioral Science, 34* 2, 180-201

Prescott, J. 1995 The evolution of competitive intelligence. *International Review of Strategic Management, 6*, 71-90

Project Management Institute 2000 A Guide to the Project Management Body of Knowledge PMBOK

Quinlan, J.R. 1989, "Unknown attribute values in induction". *6th International Workshop on Machine Learning*, anonymous eds., 164-9

Rafilson, F. 1999 The case for validity generalization. *Practical Assessment, Research & Evaluation, 2*

Ratneshwar, S., Pechmann, C. and Shocker, A.D. 1996 Goal-derived categories and the antecedents of across-category consideration. *Journal of Consumer Research, 23* 3, 240-50

Reaves, C.C. 1992 *Quantitative Research for the Behavioral Sciences*. John Wiley & Sons.

Redman, T.C. 1992 *Data Quality: Management and technology*. NY: Bantam Doubleday Dell

Reed, M. 1991 Organizations and rationality: The odd couple. *Journal of Management Studies, 28*, 559-67

Reichheld, F.F. 1993 Loyalty-based management. *Harvard Business Review, 71*, 64-73

Reichheld, F.F. and Sasser, W.E. 1990 Zero defections: Quality comes to services. *Harvard Business Review, 68* September-October , 105-11

Remenyi, D. , Sherwood-Smith, M. and White, T. 1997 *Achieving maximun value from information systems. A process approach.* Wiley

Remenyi, D. , Williams, B. , Money, A. and Swartz, E. 1998 *Doing research in business and management. An introduction to process and method.* SAGE

Roberts, M.L. 1996, "Direct marketing" in *International Encyclopedia of Business and Management*, Warner, M. eds., 1020-6

Rockart, J.F. and DeLong, D.W. 1989 *Executive support systems: The emergence of top management computer use.* Homewood, IL: Dow JonesIrwin

Roest, H. and Pieters, R. 1997 The nomological net of perceived service quality. *International Journal of Service Industry Management, 8* 4, 336-51

Rowland, G. 1995 Archetypes of systems design. *Systems practice, 8*, 277-87

Rudra, A. and Yeo, E. 1999, "Key issues in achieving data quality and consistency in data warehousing among large organizations in Australia". *Thirty second annual Hawaii international conference on system sciences*, anonymous eds.

Sakman, M.G. 1985 An empirical study of three methods of problem definition in ill-structured situations. Univeristy of Wisconsin

SAS 2005 Software Base SAS

SAS 2005 SAS/INSIGHT

SAS 2005 SAS/STAT

SAS 2005 SAS/Enterprise Miner

Sawyerr, O.O., Edbrahimi, B.P. and Thibodeaux, M.S. 2000 Executive environmental scanning, information source utilisation, and firm performance: The case of Nigeria. *Journal of Applied Management Studies, 9* 1, 95-115

Sayeed, L. and Brightman, H.J. 1994 Can information technology improve managerial problem finding? *Information & Management, 27* 6, 377-90

Schaefer, D.A. and Dillman, D.A. 1998, "Development of a standard e-mail methodology: Results of an experiment". *53rd Annual Conference of the American Association for Public Opinion Research*, anonymous eds., St. Louis, MO

Schank, R.C. 1988, "Creativity asa mechanical process" in *The nature of creativity: Contemporary psychological perspectives*, Sternberg, R.J. eds., Cambridge, UK: Cambridge University Press, 220-38

Schroder, H.M. 1989 *Managerial competence: The key to excellence.* Dubuque, Iowa: Kendal/Hunt

Schroeder, R.G. and Benbasat, I. 1975 An experimental evaluation of the relationship of uncertainty in the environment to information used by decision makers. *Decision Sciences, 6,* 556-67

Schutt, R. .K. 1996 *Investigating the social world. The process and practice of research.* Pine Forge Press

Schwab, D.P. 1999 *Research methods for organizational studies.* Lawrence Eralbaum Associates

Schwenk, C.R. 1984 Cognitive simplification processes in strategic decision-making. *Strategic Management Journal, 5* 2, 111-28

Seddon, P. 1997 A respecification and extension of the DeLone and McLean Model of IS Success. *Information Systems Research, 8* 3 , 240-53

Seddon, P. and Kiew, M.-Y. 1994, "A partial test and development of the Delong and McLean model of IS success". *International conference on information systems*, DeGross, J.I. *et* al. eds. Vancouver: 99-110

Segars, A.H. and Grover, V. 1993 Re-examining perceived ease of use and usefulness: A confirmatory factor analysis. *Management Information Systems Quarterly, 17* 4, 517-25

Senge, P.M. 1992 Mental models. *Planning Review, 20* 2, 4-10-4

Shan, N., Ziarko, W., Hamilton, H.J. and Cercone, N. 1996, "Discovering classification knowledge in databases using rough sets". *2nd. Int. Conf. Knowledge Discovery and Data Mining*, anonymous eds., 271-4

Shannon, C.E. and Weaver, W. 1949 *The mathematical theory of communication.* Urbana, IL: University of Illinois Press.

Shapiro, B.P. and Bonoma, T.V. 1984 How to segment industrial markets. *Harvard Business Review, 62* 3, 104-10

Shaver, K.G. and Scott, L.R. 1991 Person, process, and choice: The psychology of new venture creation. *Entrepreneurship Theory and Practice*, Winter, 23-42

Sherer, W.R., Maddux, J.E., Mercadante, B., Prentice-Dunn, S., Jacobs, B. and Rogers, R.W. 1982 The self-efficacy scale: Construction and validation. *Psychological Reports, 51*, 663-71

Sheth, J.N. and Parvatiyar, A. 1995 Relationship marketing in consumer markets: antecedents and consequences. *Journal of the Academy of Marketing Science, 23* 4, 255-71

Shocker, A.D., Srivastava, R.K. and Ruekert, R.W. 1994 Challenges and opportunities facing brand managment: An introduction to the special issue. *Journal Of Marketing Research, 31*, 149-58

Silver, L.S. 2000 A social-cognitivee approach to salesperson work motivation. Louisiana Tech University

Simon, H. 1955 A behavioral model of rational choice. *Quarterly Journal of Economics, 69*, 99-118

Simon, H.A. 1960 *The new science of management decision.* NY: Harper and Row

Simon, H.A. 1962, "The architecture of complexity". *American Philosophical Society*, anonymous eds., 467-82

Simon, H.A. 1965 *The shape of automation for men and management*. NY: Harper and Row

Simon, H.A. 1976 *Administrative behaviour: A study of decision making process in administrative organizations*. NY: The Free Press

Simon, H.A. 1977 *Models of discovery*. Dordrecht, Holland: D. Reidel Publishing Co.

Simon, H.A. 1978 Rationality as process and product of thought. *Journal of the American Economic Association, 68*, 1-16

Simon, H.A. and Newell, A. 1971 Human problem solving: The state of the theory in 1970. *American Psychologist, 26*, 145-59

Simoudis, E., Livezey, B. and Kerber, R. 1995, "Using Recon for data cleaning". *1st* Int. Conf. Knowledge Discovery and Data Mining, anonymous eds., AAAI Press 282-7

Sinickas, A. (2005) Cultural differences and research. Strategic Communication Management, 9 (1) 12-12

Sinkula, J.M. 1994 Market information processing and organizational learning. *Journal of Marketing, 58*, 35-45

Sinkula, J.M., Baker Jr., W.E. and Noordewier, T. 1997 A framework for market-based organizational learning: Linking values, knowledge, and behaviour. *Journal of the Academy of Marketing Science, 25* 4, 305-18

Sinkula, J.M. and Hampton, R.D. 1988 Centralization and information acquisition by in-house market research departments. *Journal of Business Research, 16* June, 337-49

Slovic, P., Fischhoff, B. and Lichtenstein, S. 1977 Behavioral decision theory. *Anual Review of Psychology, 28*, 1-39

Smith, G.F. 1989 Defining managerial problems: A framework for prescriptive theorizing. *Management Science, 35* 8, 963-81

Smith, G.F. 1990 Heuristics methods for the analysis of managerial problems. *OMEGA, 18* 6, 625-35

Smith, G.F. 1995 Classifying managerial problems: an empirical study of definitional content. *Journal of Management Studies, 32* 5, 679-706.

Soejarto, A. 2002 Business intelligence and data warehousing worldwide IT services market size and forecast, 2000-2006. Gartner

Solomon, D.J. 2001 Conducting web-based surveys. *Practical Assessment, Research & Evaluation, 7*

Song, M., Xie, J. and Dyer, B. 2000 Antecedents and consequences of marketing managers' conflict-handling behaviors. *Journal of Marketing, 64* 1, 50-66

Sorge, A. 1996, "Organizational behaviour" in *International Encyclopedia of Business and Management*, Warner, M. eds., 3793-820

Spector, P.E. 1981 *Research designs*. SAGE

Spencer, L. and Spencer, S. 1993 *Competence at work: Models for superior performance*. NY: John Wiley & Sons

Spreitzer, G., M. 1995 Psychological empowerment in the workplace: Dimensions, measurement, and validation. *Academy of Management Journal, 38* 5, 1442-65

SPSS 2003 Software SPSS 12.0 for Windows

Srinivasan, A. 1985 Alternative Measures of system Effectiveness: Associations and Implications. *Management Information Systems Quarterly, 9* 3, 243-53

Srivastava, R., Strutton, D. and Pelton, L.E. 2001 The will to win: An investigation of how sales managers can improve the quantitative aspects of their sales force's effort. *Journal of Marketing Theory and Practice, 9* 2 11

Srivastava, R.K., Shervani, T.A. and Fahey, L. 1998 Market-based assets and shareholder value: A framework for analysis. *Journal of Marketing*, *62* 1, 2-18

Stajkovic, A.D. 1998 Self-efficay and work-related task performance: A meta-analysis. *Psychological Bulletin*, *124* 2, 240-61

Stein, D.J. 1992 Schemas in the cognitive and clinical sciences. *Journal of Psychotherapy Integration*, *2*, 45-63

Stewart, D.W. and Shamdasani, P. 1998, "Focus group research" in *Handbook of applied social research methods*, Bickman, L. and Rog D. J. eds

Sticht, T.G. 1976, "Comprehending reading at work" in *Cognitive processes in comprehension*, Just, P. and Carpenter P. eds., Hillsdale, NJ: Erlbaum

Stone, D.N. 1994 Overconfidence in initial self-efficacy judgements: Effects on decision processes and performance. *Organizational Behaviour and Human Decision Processes*, *59* 3, 452-74

Straub, D.W. 1989 Validating instruments in MIS research. *Management Information Systems Quarterly*, *13* 1, 146-69

Streufert, S. 1973 Effects of information relevance on decision making in complex environments. *Memory and Cognition*, 224-8

Stubbart, C.I. 1989 Managerial cognition: A missing link in strategic management research. *The Journal of Management Studies*, *26* 4, 325-47

Suedfeld, P., Tetlock, P. and Streufert, S. 1992, "Conceptual/integrative complexity" in *Motivation and personality: Handbook of thematic content analysis*, Smith, C.P. eds., NY: Cambridge University Press, 605-11

Sullivan, D.P. and Weaver, G.R. 2000 Cultural cognition in international business research. *Management International Review*, *40* 3, 269-97

Swanson, E.B. 1987 Information channel disposition and use. *Decision Sciences*, *18* 1, 131-45

Swartzmeyer, E.G. 1987 A empirical investigation of information value : A utility approach to benefit assessment. GEORGIA STATE UNIVERSITY

Swinth, R., Gaumnitz, J. and Rodriguez, C. 1975 The decision making process: Using discrimination nets for security selection. *Decision Sciences, 6*, 439-48

Taylor, S.S. 1999 Making sense of revolutionary change: differences in members' stories. *Journal of Organizational Change Management, 12* 6 524

Thomas, K.W. and Tymon Jr., W.G. 1993 *Empowerment inventory*. Tuxedo, NY

Tversky, A. and Kahneman, D. 1972 Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207-32

Tversky, A. and Kahneman, D. 1974 Judgement under uncertainty: Heuristics and biases. *Science, 185*, 1124-31

Tversky, A. and Kahneman, D. 1981 The framing of decisions and the psychology of choice. *Science*, 211 , 453-8

Uthurusamy, R., Fayyad, U.M. and Spangler, S. 1991, "Learning useful rules from inconclusive data" in *Knowledge discovery in database*, Piatetsky-Shapiro, G. and Frawley W. J. eds. AAAI Press, 141-57

Van Birgelen, M., de Ruyter, K. and Wetzels, M. 2000 The impact of incomplete information on the use of marketing research intelligence in international service settings: An experimental study. *Journal of Service Research, 2* 4 , 372-87

Van Birgelen, M., de Ruyter, K. and Wetzels, M. 2001 What makes service research centers effective? *Journal of Service Research, 3* 3, 265-73

Vandenbosch, B. and Ginzberg, M.J. 1997 Lotus notes and collaboration: Plus ca change... *Journal of Management Information Systems, 13* 3, 65-81

Vandenbosch, B. and Higgins, C. 1996 Information acquisition and mental models: An investigation into the relationship between behavior and learning. *Information Systems Research, 7* 2, 198-214

Vandenbosch, B. and Higgins, C.A. 1995 Executive support systems and learning: A model and empirical test. *Journal of Management Information Systems*, *12* 2 99

Vandenbosch, B. and Huff, S.L. 1997 Searching and scanning: How executives obtain information from executive information systems. *MIS Quarterly*, *27* 1, 81-107

Venkatesh, V. and Davis, F.D. 2000 A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, *46* 2, 186-204

Vessey, I. 1991 Cognitive fit: A theory-based analysis of the graphs vs. tables literature. *Decision Sciences*, *22* 2

Volkema, R.J. 1983 Problem formulation in planning and design. *Management Science*, *29* 6, 639-52

Wai-yi, B.C. 1998 An information seeking and using process model in the workplace: a constructivist approach. *Asian Libraries*, *7* 12 375

Walsh, J.P. 1995 Managerial and organizational cognition: Notes from a trip down Memory Lane. *Organization Science*, *6* 3, 280-321

Walsh, J.P. and Ungson, G.R. 1991 Organizational memory. *The Academy of Management Review*, *16* 1, 57-91

Wand, Y. and Wang, R.Y. 1996 Anchoring data quality dimensions in ontological foundations. *Communications of the ACM*, *39* 11, 86-95

Wang, P. and Chan, P.S. 1995 Top management perception of strategic information processing in a turbulent environment. *Leadership & Organization Development Journal*, *16* 7, 33-43

Wang, P. and Spiegel, T. 1994 Database marketing and its measurements of success: Designing a managerial instrument to calculate the value of a repeat customer base. *Journal of Direct Marketing*, *8* 2, 73-81

Wang, R.Y., Reddy, M.P. and Kon, H.B. 1995 Toward quality data: An attribute-based approach. *Decision Support Systems*, *13* 3/4, 349-72

Wang, R.Y., Storey, V.C. and Firth, C. 1995 A framework for analysis of data quality research. *IEEE Transactions on Knowledge and Data Engineering*, *7* 4, 349-72

Wang, R.Y. and Strong, D.M. 1996 Beyond accuracy: what data quality means to data consumers. *Journal of Management Information Systems*, *12* 4, 5-34

Wanzenried, J.W. and Powell, F.C. 1993 Source credibility and dimensional stability: A test of the Leathers personal credibility scale using perceptions of three presidential candidates. *Perceptual and Motor Skills*, *77*, 403-6

Wardley, M. and R. Blumstein 2004 Worldwide CRM applications 2004-2008 forecast update and 2003 Vendor Shares. IDC

Weick, K. 1979 *The social psychology of organizing*. WCB/McGraw-Hill

Weick, K.E. 1976 Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, *21* 1, 1-19

Weick, K.E. 1993 Collective mind in organizations: Heedful interrelating on flight decks. *Administrative Science Quarterly*, *38* 3, 357-81

Weick, K.E. 1993 The collapse of sensemaking in organizations: The Mann Gulch disaster. *Administrative Science Quarterly*, *38* 4, 628-52

Weick, K.E. 1995 *Sensemaking in organizations*. SAGE

Weick, K.E. 2000 *Making sense of the organization*. Blackwell Publishers

Wetherbe, J.C. 1991 Executive information requirements: Getting it right. *Management Information Systems Quarterly*, *15* 1, 50-65

Wierenga, B. and Van Bruggen, G.H. 1998 The dependant variable in research into the effects of creativity support systems: Quality and quantity of ideas. *Management Information Systems Quarterly*, *22* 1, 81-7

Winkler, R.L. 1982 Research directions in decision making under uncertainty. *Decision Sciences, 13* 4, 517-33

Witteloostuijn, A.V. 1996, "Contexts and environments" in *International Encyclopedia of Business and Management*, Warner eds., 752-61

Wixom, B.H. and Watson, H.J. 2001 An empirical investigation of the factors affecting data warehousing success. *Management Information Systems Quarterly, 25* 1, 17-41

Wolf, K.N. 1997 Predicting positive self-efficacy in group problem solving. *Human Resource Development Quarterly, 8* 2 , 155-69

Wood, R. 1986 Task complexity: Definition of the construct. *Organizational Behaviour and Human Decision Processes, 37*, 60-82

Wood, R. and Bandura, A. 1989 Impact of conceptions of ability on self-regulatory mechanisms and complex decision-making. *Journal of Personality and Social Psychology, 56* 3, 407-15

Wood, R. and Bandura, A. 1989 Social cognitive theory of organizational management. *Academy of Management Review, 14* 3, 361-84

Woodruff, S.L. and Cashman, J.F. 1993 Task, domain, and general efficacy: A re-examination of the self-efficacy scale. *Psychological Reports, 72*, 423-32

Wright, W.F. 1980 Cognitive information processing biases: Implications for producers and users of financial information. *Decision Sciences, 11* 2

Yeoh, P.-L. 2000 Information acquisition activities: A study of global start-up exporting companies. *Journal of International Marketing, 8* 3, 36-60

Yung, Y.-F. and Bentler, P.M. 1996, "Bootstrapping techniques in analysis of mean and covariance structures." in *Advanced structural equation modeling: Issues and techniques*, Marcoulides, G.A. and Schumacker R. E. eds., Mahwah, NJ: Lawrence Eralbaum Associates, 195-226

Zhu, Z. (1999) The practice of multimodal approaches, the challenge of cross-cultural communication, and the search for responses. Human Relations, 52 (5), 579-607

Zmud, R.W. 1978 An empirical investigation of the dimensionality of the concept of information. *Decision Sciences*, *9* 2, 187-95

APPENDIX A: CONCEPT VALIDATION RESEARCH FOCUS GROUP

Description

Session Title: Data warehouses in production. Focus group

Context: Partners 2000 conference. Orlando (FL)

Date, time, place: Tuesday, September 26, 2000, 4:30-5:50 P.M., Europe 6 (Dolphin Hotel)

Facilitator/Moderator: Raul M. Abril

Participants: Retail industry: 1 (Denmark)

Financial industry: 1 (Israel), 1 (Netherlands), 1(Argentina)

Telecommunications industry: 1(Spain), 1-fix line- (Austria), 1-cellular- (Austria)

Method

Potential barriers to open communication (e.g. competition) were avoided by limiting the group to one participant per industry and country. The fix line provider and the cellular provider from Austria did not regard each other as a competitor. Two invited retail firms were not able to attend. The qualification criteria for participating were: (1) More than one year of data warehouse in production, (2) Responsibility for the usage of the data warehouse either as a user or as a service function. Five questions were e-mailed in advance to the participants. Clarifications were offered over the phone. Answers were provided in a round table discussion with open discussion after each question/round. The questions were:

Q1: What type of measures do you have for the value contribution of your data warehouse?

Q2: What measures would you recommend for the value contribution of your data wrehouse?

Q3: What kind of barriers do you find in promoting usage? (e.g. data quality)

Q4: What type of queries do you have?

The facilitator had a questionnaire ready for recording the answers. The group answers were e-mailed one week later to the participants asking for confirmation. After active follow up we had 3 confirmations out of seven without changing the initial answers. We did not have any (dis)confirmation from the other four.

Focus group results

1. Most of the participants have measures for the value contribution of their data warehouses. Number of supported functions
2. Data quality and organizational financial performance are the measures recommended.
3. Organizational financial performance measures were recommended by the majority (five) of the participants for evaluating a data warehouse although participants did not have them.
4. Data quality was recommended by three of the participants for evaluating a data warehouse.
5. The definition "Manager's perception on the extent to which the information derived from the available customer relationship data in the DW fits with the manager's customer relationship management needs for problem understanding" of the latent construct quality on DW customer relationship data for problem enactment (QUALITY) was considered clear by the participants. I slightly changed some of the terms in order to make this definition more understandable to the participants.
6. The majority (five) of the participants considered information management competence as a barrier promoting usage of the information derived from their DW.
7. Knowledge about the business process (e.g. CRM) was considered by three participants to be the second barrier to promoting usage of the information derived from their data warehouse.
8. Only one participant considered that user's job function is not a moderator of usage of the information derived from their data warehouses.
9. Most of the participants estimated that (1) between 75% and 90% of the queries are of a "What happened" nature (2) between 5% and 20% of the queries are of a "Why did it happen?" nature, and (3) between 0% and 5% of the queries are of a "What will happen?" nature. Two participants reported that they did not have any queries of a "What will happen?" nature.

APPENDIX B: CONTENT VALIDITY RESEARCH FOCUS GROUP. QUESTIONNAIRE.v1

Description

Title: Scholars focus group for data warehousing and CRM

Date, place: February-March 2003, Virtual

Facilitator/Moderator: Raul M. Abril

Participants (8):

RESPONSE STATISTICS				
Delivered	8	32.00%	Response rate	32.00%
WIP (delivered 1st round)	0	0.00%		
Dropped (delivered 1sr round)	0	0.00%		
Accepted (but did not participate)	2	8.00%		
Not Readed	0	0.00%		
Declined	8	32.00%		
Bounced	0	0.00%		
Received (no response)	7	28.00%		
Deleted w/o response	0	0.00%		
Total	25	100.00%		

Prof. "Barna" (he teaches in an international university in Spain. He provided expert opinion on data quality scales)

Prof. "Mallorca" (one of the top experts in self-efficacy and close collaborator of Prof. A. Bandura, teaching in an US university. He provided expert opinion on the self-efficacy scale)

Prof. K. Wolf (She teaches at the School of Allied Medicine, The Ohio State University. She provided input on the changes I did on her self-efficacy scale and also commented the cognitive scales)

Prof "Pearson" (he teaches in one of the top non-US Business Schools. He focused on data quality scales)

Prof. R. Hackney (he teaches at the Business School, Manchester Metropolitan University. He provided comments from a qualitative/interpretative posture with focus in behavioral and data quality scales)

Prof. "Dublin" (she teaches at a major Irish university. She focused on behavioral scales)

Prof. M. Jeffery (He teaches at Kellogg School of Management, Northwestern University. He provided input about all the scales)

Prof. "Duke" (he teaches at a top US-Business School. He is an expert in CRM analytics. He provided input about all the scales)

Method

The profile of the participants had to satisfy the following requirements (i) to be a scholar, (ii) be knowledgeable in at least one of the constructs in my research. Participants were identified taking into consideration their research contributions to the bodies of knowledge involved in my research model. I targeted 25 scholars and ended with 8 participants. Invitations were e-mailed to targeted participants (see the template used at the end of this appendix) together with the highlights of this research. A Delphi technique via e-mail was used taking two rounds of questions (see the templates used at the end of this appendix). Participants were updated with the aggregated findings of the first round in the second round. Endless follow up phone calls were required. Most of the answers were received by e-mail and a few over the phone.

Focus group results

Weaknesses

W1: The Self-efficacy items #50 to #53 (low difficulty level) were not consistent with the moderate and high difficulty subscales in their task domain.

W2: Most of the items were positively loaded which is likely to induce artificially high reliability.

W3: Response format for scales in Part III (rarely, sometimes, frequently...) are too vague and will have different meaning for different people, rendering the data useless.

Risks

R1: Several constructs with very few items

R2: A positivistic approach is a disaster unless coupled with more qualitative methods

R3: Questionnaire is very long

INVITATION TO THE CONTENT VALIDITY RESEARCH FOCUS GROUP

Dear Field1,

As a research associate at Henley Management College, UK, I am conducting a research project on Data Warehousing and Customer Relationship Management. The objective of the research is to better understand selected customer relationship issues and their impact on the use of data warehouses. Professor Robert M. O'Keefe (Head of School, School of Management, University of Surrey, UK), professor Joe F. Hair, Jr. (Director, Entrepreneurship Institute, Louisiana State University, USA), and professor David Price (Director of Studies, Doctoral Programmes, Henley Management College, UK) are my supervisors on the project.

I am writing to you because I am looking for members for a Scholars Focus Group to review the operationalization of my research model. I kindly invite you to participate in this Scholars Focus Group. The bellow attached MS Powerpoint file contains a presentation to this research.

I will use the Delphi technique via e-mail. I anticipate only a few e-mail exchanges. Your identity will be kept in complete confidence, and eventually will only be reported in a composite form with those of other respondents (i.e. no references will be made to specific individuals or names of institutions in future potential reports). Henley Management College has the highest standards of integrity to be adhered to by those undertaking research. This research is not an exception and it is under the rule of Henley's Policy on Research Practice.

Field1, it would be an honor for me to have your input. Pls., do not hesitate in asking me for clarifications. I am looking forward your positive answer.

Thank you in advance and best regards,

RA

<Presentation PPT>

CONTENT VALIDITY RESEARCH FOCUS GROUP. ROUND #1.

Dear Field1,

Thank you again for your willingness and interest in working with the Scholars Focus Group for **Data Warehousing and Customer Relationship Management**. Bellow you have the **questionnaire to be reviewed** by you (pls., do not fulfill it).

<Questionnaire .doc>--

Bellow you have **my questions (Round #1) about the above questionnaire**:

- valuate the scales (i.e. the scales measure what they are supposed to measure)
- have the definitions of constructs been done properly?
- does each scale cover all the dimensionalities of its construct?
- is there any contamination? (i.e. scale capturing characteristics not specifically included in the definition of its construct)
- overall, do the scales seem to have content validity?

Please, e-mail your feedback to me latest by xxxxxx. Let me know if that is impossible. Let me also know if you have problems opening the attachments. The attachments are virus checked with the January 11, '03 version v4.5.1 SP1of McAfee. Once again thank you very much for your help. I look forward to receive your feedback.

Best regards,

Reference material:

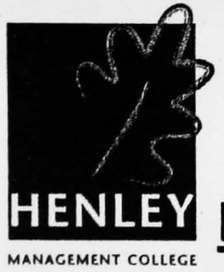
#1 Definitions of constructs (i.e. Part I, II, and III in the questionnaire)

<def part I .doc><def part II.doc><def part III.doc>

#2 The highlights of this research.

<Presentation PPT>

QUESTIONNAIRE.v1



Data Warehousing and Customer Relationship Management

First, thank you very much for your help. By answering this questionnaire, you can help marketers to more effectively execute CRM tasks. The questions focus on the quality of data in your firm's data warehouse (i.e. integrated collection of customer data from various sources) and ways to enhance the benefits of CRM activities. Hopefully, you can easily answer the questions based on your experience in dealing with customer relationship issues. Based on our pretest, it should take no more than 20 minutes to complete the survey.
There are not right or wrong answers. Your responses will be kept in complete confidence and only be reported in a composite form with those of other respondents (i.e. no references will be made to specific individuals or names of organizations).
If you have any questions or concerns please contact me using the e-mail address listed below.

Best regards,

Raul M. Abril
Research Associate
Henley Management College, UK
r.abril@cox.net

N. B.: Raul is following a doctoral program at Henley Management College, UK. His research focuses on the relationships between managerial problem solving, information search behavior in customer relationship management processes supported by a data warehouse, and data quality. As a practitioner, he has been working in the area of decision support systems for eighteen years.

DETAILS ABOUT YOU AND YOUR CRM WORK ENVIRONMENT

The first 9 questions relate to your experiences in a CRM work environment. Please respond using the scale to the right of each question.

	Less than 6 months	Between 6 months and 1 year	Between 1 and 3 years	Between 3 and 5 years	More than 5 years
1. How long have you been involved in customer relationship management tasks?	○	○	○	○	○
2. How long have you been using your **company's data warehouse customer relationship data** to help you understand issues?	○	○	○	○	○
3. **How long** has your data warehouse been supporting CRM functions (any of them)?	○	○	○	○	○

	No time at all 0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	All of my time 100%
4. In a **typical week** how much time do you spend **understanding CRM issues** (e.g. challenges, problems)	○	○	○	○	○	○	○	○	○	○	○
5. In a **typical week** how much time do you spend **searching/requesting customer information** from your firm's data warehouse?	○	○	○	○	○	○	○	○	○	○	○

	Sales support	Marketing support	Customer service support	Other CRM functions
6. Which CRM functions are supported by your data warehouse? Check all that apply.	☐	☐	☐	☐

Sales Support = sales analysis, sales history, contact management, lead tracking, etc.
Marketing Support = customer analysis, channel analysis, channel management, campaign management, etc.
Customer Service Support = Call management, problem tracking, customer history, etc.

	Sales support	Marketing support	Customer service support	Other CRM functions
7. What is the **primary** CRM function of **your personal** job?.	☐	☐	☐	☐

To what extent do you agree are disagree with the following statements.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
8. My firm's data warehouse is an **important source** of customer relationship data for my personal job.	☐	☐	☐	☐	☐
9. My firm's CRM activities target **individuals** as customers (i.e., not just other businesses or channel organizations)	☐	☐	☐	☐	☐

PART I: YOUR FIRM'S DATA WAREHOUSE

Below there are 8 statements about the **actual** contents of your data warehouse in terms of customer relationship types of data. For each, please indicate the extent to which you feel it reflects the situation in your organization. Click on a number on the scale to identify the degree to which you agree or disagree with the statement. **To what extent do you agree or disagree that your firm's data warehouse includes the following?**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
10. **customer's motivational data** (e.g.needs)	○	○	○	○	○	○	○
11. **customer's attitudinal data** (e.g. trust, satisfaction)	○	○	!!! Customer Insights Items #10 to #15				○
12. **relationship measures** (e.g. attrition risk, life-time value)	○	○					○
13. **customer relationship events** (e.g., transactions, contacts, complaints)	○	○	○	○	○	○	○
14. **customer's geodemographic data** (e.g., age, gender)	○	○	○	○	○	○	○
15. **customer's behavioral data** (e.g., dominant purchase tendency, dominant channel usage habits)	○	○	○	○	○	○	○
16. **your competitors** (e.g., products, campaigns, channels)	○	○	!!! Competitor Insights Items #16 and #17				○
17. **opportunities for competitive advantage**	○	○	○	○	○	○	○

Remember, we focus on *understanding* (i.e., not in *solving*). Below are two questions about the usefulness of your data warehouse's customer relationship data for <u>understanding</u> customer relationship problems. **Please, give your opinion on ....**

	Not at all	A little extent	Some extent	Large extent	Very relevant
18. How **relevant** for understanding customer relationship problems is the customer relationship data in your firm's data warehouse?	○	!!! Customer relationship problem enactment Items #18 to #21			

	Never	Rarely	Sometimes	Frequently	Most of the time
19. **How often is** the data you get from your data warehouse exactly what you require to understand customer relationship problems?	○	○	○	○	○

Below there are 20 statements about some characteristics of the customer relationship data in your data warehouse. Please respond as it describes the situation in your firm. Click on a number on the scale to identify the degree to which you agree or disagree with the statement. **Your firm's data warehouse customer relationship data . . .**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
20. enables you **to find out the relevant factors** (i.e., components) of customer relationship problems.	○	○	○	○	○	○	○
21. enables you **to find out the relevant relationships between the factors** (i.e., components) in customer relationship problems.	○	○	○	○	○	○	○
22. enables you to **easily identify the number of customers** your organization has.	○	○	○	○	○	○	○
23. integrates historic and operational data **at the customer level**.	○	○	○	○	○	○	○
24. allows **a consistent 360° view** of each customer.	○	○	○	○	○	○	○
25. **effectively integrates** customer relationship data from systems that serve **different functional areas** (e.g. sales, customer service).	○	○	○	○	○	○	○
26. **effectively integrates** customer relationship data from **a variety of data sources**.	○	○	○	○	○	○	○
27. **can be interpreted** in several ways that **lead** to different but acceptable customer relationship **problem formulations**.	○	○	○	○	○	○	○

!!! Data Integration Items #22 to #26

!!! Data Equivocality Items #27, #29, #36

Your firm's data warehouse customer relationship data . . . .

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
28. includes incorrect **data.**	○	○	○	○	○	○	○
29. **means** different things to different people.	○	○	○	○	○	○	○
30. can **easily be summarized** at different levels.	○	○	○	○	○	○	○
31. includes **missing data.**	○	○	!!! Data Richness Items #30, #32, #33, #34, #37				
32. **is organized** in a meaningful way.	○	○					
33. includes **valuable information** for understanding customer relationship problems.	○	○	○	○	○	○	○
34. **provides new/unanticipated insights** for understanding customer relationship problems.	○	○	○	○	○	○	○
35. is a **trusted source** for understanding customer relationship problems.	○	○	○	○	○	○	○
36. can support **more than one plausible formulation** for the customer relationship **problems** you face.	○	○	!!! Data Trustworthiness Items #28, #31, #35, #38, #39				
37. challenges **a-priori interpretations** of customer relationship problems.	○	○	○	○	○	○	○
38. is **accurate.**	○	○	○	○	○	○	○
39. does **not require double-checks** with other sources.	○	○	○	○	○	○	○

PART II: UNDERSTANDING CUSTOMER RELATIONSHIP PROBLEMS

Below there are 10 statements that relate to your understanding of customer relationship issues. Please indicate the extent to which each statement reflects the situation in your personal case.

Think about the customer relationship problems you have faced in the last 6 months. **To what extent do you agree or disagree that the use of data warehouse customer relationship data has personally impacted your understanding of customer relationship problems?**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
40. I have a **clearer sense** of customer relationship problems (i.e., better focus).	○	○	○	○	○	○	○
41. I reach **sharper final interpretations** of customer relationship problems.	○	○	○	○	○	○	○
42. I have a **better understanding** of important customer relationship trends.	○	○	○	○	○	○	○
43. I have **better insights** into the customer relationship threats facing my firm.	○	○	○	○	○	○	○
44. **I look for more information** in understanding customer relationship problems.	○	○	○	○	○	○	○
45. **I analyze more** relevant information before formulating a plausible customer relationship problem statement.	○	○	○	○	○	○	○
46. Quantitative analytic techniques **are more important to me** in understanding a customer relationship problem.	○	○	○	○	○	○	○

!!! Customer Relationship Problem Declarative Knowledge Enhancement Items #40 to #43

!!! Customer Relationship Problem Enactment Procedural Knowledge Enhancement Items #44 to #47

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
47. I am **more effective focusing** my attention on crucial information.	○	○	○	○	○	○	○
48. My final interpretations of customer relationship problems have **more factors** (i.e., problem components)	○	○	○	○	○	○	○
49. I reach final interpretations of customer relationship problems with **more interconnections among the factors**.	○	○					

!!! Customer Relationship Problem Integrative Complexity Enhancement Items #48 and #49

Some situations can make it harder or easier to define a plausible customer relationship problem. Below there are 12 statements relating to your degree of confidence in a number of tasks. **Please rate your degree of confidence in your ability to complete the following tasks.**

	No chance 0	1	2	3	4	A 50-50 chance 5	6	7	8	9	Completely certain 10
50. I can recognize **familiar** customer relationship problems.	○	○	○	○	○	○	○	○	○	○	○
51. I can formulate a plausible customer relationship problem statement when I **have access to the needed data**.	○	○									
52. I can recognize **easily answerable** questions.	○	○	○	○	○	○	○	○	○	○	○
53. I can confirm conclusions for **simple problems** looking at customer relationship data.	○	○	○	○	○	○	○	○	○	○	○

!!! Customer Relationship Problem Enacting Self-Efficacy Items #50 to #61
Low difficulty: Items #50 to #53
Moderate difficulty: Items #54 to #57
High difficulty: Items #58 to #61

	No chance 0	1	2	3	4	A 50-50 chance 5	6	7	8	9	Completely certain 10
54. I can identify **new factors** contributing to a customer relationship problem.	○	○	○	○	○	○	○	○	○	○	○
55. I can reach a plausible customer relationship problem statement under **moderate time pressure**.	○	○	○	○	○	○	○	○	○	○	○
56. I can identify **new questions** whose answers might give new insights.	○	○	○	○	○	○	○	○	○	○	○
57. I can explain cause-effect relationships in a problem when I have **limited information**.	○	○	○	○	○	○	○	○	○	○	○
58. I can understand **radically new types** of customer relationship problems.	○	○	○	○	○	○	○	○	○	○	○
59. I can establish a plausible statement about a **high stakes** (e.g., risk of losing a market segment) customer relationship problem.	○	○	○	○	○	○	○	○	○	○	○
60. I can explore **different patterns or trends** in customer relationship data.	○	○	○	○	○	○	○	○	○	○	○
61. I can analyze customer relationship data in a situation where there are **competing goals/objectives** about the desired customer relationship status (e.g., keep or terminate customer relationship).	○	○	○	○	○	○	○	○	○	○	○

PART III: INFORMATION SEARCH/REQUEST BEHAVIOR

Below are 5 statements about your way of searching/requesting information. Please indicate how frequently you search for/request information.

Think of the situations in the last 6 months where you have been engaged in searching/requesting customer relationship information.

	Never	Rarely	Sometimes	Frequently	Most of the time
62. **How often** do you search/request information on **broad trend information** in your firm's data warehouse?	○	○	○	○	○
63. **How often** do you search/request information on **unknown problems/key issues** in your firm's data warehouse?	○	○			
64. **How often** do you search/request information on **new ideas and strategies** in your firm's data warehouse?	○	○	○	○	○
65. **How often** do you use your data warehouse to search/request information to help you to understand **specific customer problems** or for **specific account information**?	○	○	○	○	○
66. **How often** do you use your data warehouse to search/request information to help you **react** to a problem?	○	○			

!!! Intensity of data warehouse scanning
Items #62 to #64 and #67

!!! Intensity of data warehouse focused search
Items #65, #66 and #68

Listed below are 2 statements about the percentage of your time you search/request information in either an exploratory way or a focused way. For these questions, please use the following definitions.

Scanning means proactive browsing through information **without a particular problem** to solve (e.g., routine performance monitoring of a campaign). **Focused** search/request means reactive looking for information **specific to a problem** to be addressed or question to be answered (e.g., information regarding a dissatisfied customer).

Please indicate the proportion of your time in a **typical week** you spend in the two types of searching shown below. The total of the two should be 100% or less.

67. on data warehouse **scanning**. _____
68. on data warehouse **focused** search/request. _____
Total 100%

!!! **Intensity of data warehouse scanning**
Items #62 to #64 and #67

!!! **Intensity of data warehouse focused search**
Items #65, #66 and #68

Think of the 'rules of thumb' that you have applied in the last 6 months in order to simplify your search/request for information when you need to understand a customer relationship problem. **How often do you search/request data warehouse customer information using the following strategies?**

	Never	Rarely	Sometimes	Frequently	Most of the time
69. **"The best 'problem template' at hand"**: Looking for **common information,** disregarding differences, between the situation that you have to understand and one or several **'problem template(s)'** that you have at hand.	○	○	○	○	
70. **"The test"**: Looking for information in order to identify an acceptable **sample/target** in order to be able to generalize to a larger scale.	○	○	○	○	○
71. **"The last occurrence"**: Looking for information about **recent relevant precedent** situations.					
72. **"The 'canned' queries"**: Using **predefined** questions.					
73. **"The trial and error"**: **Adjusting** your next question with the information that you got in a previous question.					
74. **"The divide & conquer"**: **Adding** successive selection **criteria** to your initial question.					
75 **"The confirmatory"**: Looking for information that **confirms** your belief about the problem.					
76. **"The diagnostic"**: Looking for information that **confirms** known symptoms.					

!!! Intensity of data warehouse representativeness search heuristic Items #69, #70, and #77

!!! Intensity of data warehouse availability search heuristic Items #71, #72, and #78

!!! Intensity of data warehouse anchoring and adjustment search heuristic Items #73, #74, and #79

!!! Intensity of data warehouse positivy search Heuristic Items #75, #76, and #80

Please indicate the proportion of your time in a **typical week** you spend (the total should be 100% or less).

77. **Using** "The best 'problem template' at hand" **and/or** "The test" search/request strategies

78. **Using** "The last occurrence" **and/or** "The 'canned' queries" search/request strategies

79. **Using** "The trial and error" **and/or** "The divide and conquer" search/request strategies

80. **Using** "The confirmatory" **and/or** "The diagnostic" search/request strategies

Total 100%

!!! Intensity of data warehouse representativeness search heuristic Items #69, #70, and #77

!!! Intensity of data warehouse availability search heuristic Items #71, #72, and #78

!!! Intensity of data warehouse anchoring and adjustment search heuristic Items #73, #74, and #79

!!! Intensity of data warehouse positivy search Heuristic Items #75, #76, and #80

RESPONDENT'S PERSONAL DATA

The following information will only be used to provide an overall picture of the total sample. All individual or company specific data will be retained in the strictest confidence. ***All items marked with an asterisk are optional and respondents will only be followed up personally when they have indicated a willingness to do so.***

Name

Position Held

Function

- O Bus. Development
- O Customer Service (e.g. Call Center)
- O E-Business
- O Executive Office
- O Fraud Management
- O Marketing - General
- O Marketing – CRM (e.g. Research, Channel, Campaign, Segment)
- O Merchandising
- O Sales (e.g. Region, Account)
- O Strategic Planning
- O Other (Please specify)

Level*

- O Board Member (Chairman, VC)
- O Business Head (CEO, President)
- O Bus Function Leader (CIO, CMO)
- O Sr. Dept Head (Director, AVP, VP)
- O Other Dept/Function Head (Manager)
- O Project/Team Leader (Project Mgr)
- O Analyst
- O Individual Contributor (e.g. Representative)
- O Other (Please specify)

Organization / Company

What is your data warehouse' s database? *

- ○ DB2
- ○ Oracle
- ○ SQL-Server
- ○ Teradata
- ○ Other (Please specify) _____

What business(es) is your company in? *

- ○ Manufacturing
- ○ Retail
- ○ Financial/Insurance
- ○ Telecommunications
- ○ Pharmaceuticals
- ○ Chemical
- ○ Computer & IT
- ○ Other (Please specify) _____

Would you like to receive a summary of the research findings? ○ Yes ○ No

If "YES", e-mail address:

Telephone contact _____

e-mail address _____

THANK YOU, FOR TAKING THE TIME TO RESPOND TO THIS SURVEY.

Please send any comments to r.abril@cox.net

Date: Thursday, 10/11/2001 7:06 PM MDT

Submit

CONTENT VALIDITY RESEARCH FOCUS GROUP. ROUND #2.

Dear Field1,

Thank you again for your willingness and interest in working with this Scholars Focus Group for **Data Warehousing and Customer Relationship Management**. Attached, please, find the aggregated feedback (weaknesses and risks) of this Focus group in relation to my questions in round #1.

Bellow you have my questions **(Round #2 and last)** about the attached questionnaire:

- comment on whether the weaknesses identified have been addressed in this new version of the questionnaire
- comment on further explanations or information to include
- comment the sequence of questions.
- regarding the informants' response qualification criteria (see slide #12 in research highlights): is it appropriate?
- comment the cover letter in the questionnaire
- comment the Nota Benne at the bottom of the cover letter in the questionnaire (i.e. disclosing that I am currently a practitioner). Furthermore, do you support that I do NOT disclose the name of my employer?
- comment the personal data items at the end of the questionnaire.
- Any further comment

Field1, please, e-mail your feedback to me latest by xxxxxxx. Let me know if that is impossible. Let me also know if you have problems opening the attachments. The attachments are virus checked with the January 11, '03 version v4.5.1 SP1of McAfee. Once again thank you very much for your help. I look forward to receive your feedback.

Best regards,

RA

Aggregated feedback (weaknesses and risks) of this Focus group in relation to my questions in round #1. Questions, assessments, and recommendations are not included

Weaknesses

W1: The Self-efficacy items #50 to #53 (Low difficulty level) are not consistent with the moderate and high difficulty subscales in their task domain.

W2: Most of the items are positively loaded which is likely to induce artificially high reliability

W3: Response format for scales in Part III (rarely, sometimes, frequently...) are too vague and will have different meaning for different people, rendering the data useless.

Risks

R1: Several constructs with very few items

R2: A positivistic approach is a disaster unless coupled with more qualitative methods

R3: Questionnaire is very long

APPENDIX C: FACE VALIDITY RESEARCH FOCUS GROUP

Description

Title: Practitioners focus group for data warehousing and CRM

Date, place: April-May 2003, Virtual

Facilitator/Moderator: Raul M. Abril

Participants (11):

RESPONSE STATISTICS	at June 23th, 2003		<--I cut here (Pilot starts)	
Delivered	6	15.79%	Response rate	28.95%
WIP (delivered 1st round)	4	10.53%		
Dropped (delivered 1sr round)	1	2.63%		
Accepted	4	10.53%	<--Still opportunity	
Not Readed	3	7.89%		
Declined	7	18.42%		
Bounced	4	10.53%		
Received (no response)	8	21.05%		
Deleted w/o response	1	2.63%		
Total	38	100.00%		

Mr. "Thor", CRM Practitioner based in Scandinavia

Mr. "Dali", IT practitioner based in Iberia

Dr. "Arthur", MIS Practitioner based in USA, before in UK

Mr. "Sabbath" CRM practitioner with international experience based in US

Mr. "Big Eye" , CRM market analyst based in US

Mr. "Canadian Banker", CRM practitioner in major Canadian bank based in Canada

Mr. "Telephone", CRM practitioner in major Telco firm based in US

Mr. "IQ" Information quality practitioner based in US

Dr. "DBA" MIS Practitioner based in Scandinavia

Dr. "Bill Gates", MIS Practitioner based in UK

Dr. "DaVinci", Marketing senior executive with international experience Based in Italy

Method

The profile of the participants had to satisfy the following requirements (i) to be a practitioner in CRM/Marketing and/or IT/MIS, (ii) more than five years of experience in CRM/Marketing and/or IT/MIS, and (iii) overall, to have an international representation. Participants were identified taking into consideration my personal past/actual network of professional contacts. I targeted 38 practitioners and ended with 6 participants delivering responses in the two rounds and five delivering responses only in the first round.

Invitations were e-mailed to targeted participants (see the template used at the end of this appendix). A Delphi technique via e-mail/phone was used taking two rounds of questions (see the templates used at the end of this appendix). Participants were updated with the aggregated findings of the first round in the second round. The amount of effort following up the participants was huge and unexpected. Most of the answers were received by e-mail and a few over the phone.

The first contact e-mail combined the invitation and the first round of questions. This first round achieved 11 responses (28.95% response rate). The first round of questions was designed in order to engage participants in the focus group (i.e. maximize response rate). The expected amount of time required by the participants for the first round was not too big and the type of questions, although very important for me, they were 'easy'. The second round was very demanding for the participants in terms of time and effort and this resulted in a drop of the response rate. One participant openly said 'too much, I'am sorry' and five never responded to my follow ups. Six participants completed the second round largely exceeding my expectations with the thoroughness of their comments.

Focus group results

Weaknesses

W1: The HTML form is too wide forcing to keep scrolling back and forth.

W2: Percentage is not a good scale (see items #4 and #5).

W3: 7-point scale is too much.

W4: Once a response is provided you cannot leave it blank (at the item level).

W5: "No responses" are not allowed (i.e. the questionnaire forces for a response before submitting) –two participants-.

W6: Too many different scales (two participants).

W7: Item #14 geodemographic data is not age gender, etc. it is more of income level based on neighborhoods.

W8: Item #18 the scale is a little strange.

W9: Item #19 the scale is a little strange.

W10: Item #23 should include a reference to "data timeliness".

W11: The cover letter does not inform about the sections in the questionnaire.

W12: The cover letter does not provide instructions for fulfilling/submitting.

Risks

R1: Questionnaire is very long.

R2: Two URL's makes it complex.

R3: PDF format is perceived riskier (HTML gets 9 votes and PDF 2 votes).

R4: Item #24 is misleading in the negative.

R5: Item #25 "partially" dilutes the meaning of the question..

R6: Item #28 "Includes incorrect data" - again that's a negative.

R7: Item #44 - look where?.

R8: The sequence of items #65, #66, #67 and #68 might confuse informants.

R9: The cover letter (i) does not address the research problem and (ii) in the first place want some assistance from the respondents to supply empirical data. "To help marketers", I couldn't care less about, but to help a serious hardworking doctorate associate yes, I would perhaps give 20 minutes of my time filling in the questionnaire!".

R10: - Avoid splitting scales (sections) of the questionnaire, which need to be dealt with by the respondent in one go!

R11: Q starts with what I would call 'demographic' data.

R12: Copyrighted statements can make respondents drop filling in the Q, because they don't want to be contributing to others commercial business!!!!.

R13: - If these scales are coming from other authors are you in your right to even mention copyright.

INVITATION TO THE FACE VALIDITY RESEARCH FOCUS GROUP AND ROUND #1

Dear Field1

As a research associate at Henley Management College, www.henleymc.ac.uk , I am conducting an academic non-profit research project on Data Warehousing and Customer Relationship Management. The objectives of the research are to better understand

- CRM managers' perceptions on the quality of data warehouse' customer relationship data,
- the impact of data warehouse' customer relationship data in CRM managers' ability understanding CRM problems –e.g. identification of the components of a customer relationship problem statement-, and
- the impact of data warehouse' customer relationship data in CRM managers' information search behaviour –e.g. search strategies-.

The practical value of this research will be:

- Results about objective #1 are important because It might be a gap between CRM managers' perceptions and the 'hard' technical metrics that the IS/IT functional owners have. Also, such results could be used in a data warehouse scorecard
- Results about objective #2 are important because they will constitute tangible evidence of contribution of the data warehouse to the organization
- Results about objective #3 are important because they can be considered in adjusting/fine tuning the available query tools to the benefit of CRM managers

Professor Robert M. O'Keefe (Head of School, School of Management, University of Surrey, UK), professor Joe F. Hair, Jr. (Director, Entrepreneurship Institute, Louisiana State University, USA), and professor David Price (Director of Studies, Doctoral Programmes, Henley Management College, UK) are my supervisors on the project.

I am seeking knowledgeable people in the industry such as yourself to provide feedback on a questionnaire to be used in my research. **Please, notice that I am NOT asking you for fulfilling any questionnaire but for your critical opinion about it.** I kindly invite you to participate in a Practitioners Focus Group which will be instrumental in collecting your feedback.

I anticipate only a couple of e-mail exchanges. Your identity will be kept in complete confidence, and eventually will only be reported in a composite form with those of other respondents (i.e. no references will be made to specific individuals or names of institutions in future potential reports). Henley Management College has the highest standards of integrity to be adhered to by those undertaking research. This research is not an exception and it is under the rule of Henley's Policy on Research Practice.

Field1, it would be an honor for me to have your input. I am looking forward your feedback. Bellow you have my questions (Round #1) about the questionnaire (two formats):

- Valuate the htm format versus the pdf format

- **What is the best format? Why?**

- Valuate the htm format (please, do not consider that the form is too wide!!!. I will fix it)

- **Is the lay-out okay? (e.g. spacing, font)**
- **Any issues opening/looking at this format?**

- Valuate the pdf format

- **Is the lay-out okay? (e.g. spacing, font)**
- **Any issues opening/looking at this format?**

These are the URLs that I would attach in an email to candidate informants

---

http://www.henleymc.ac.uk/quest/datawarehousing.pdf
http://www.henleymc.ac.uk/quest/40803.htm

Please, e-mail your feedback to me latest by xxxxxxx. Let me know if that is impossible. Once again thank you very much for your help.

Best regards,

Raul M. Abril
HMC Research Associate
ra.dba15@henleymc.ac.uk

FACE VALIDITY RESEARCH FOCUS GROUP. ROUND #2.

Dear Field1,

Thank you again for your willingness and interest in working with this Practitioners Focus Group for **Data Warehousing and Customer Relationship Management**. Bellow you have my questions (Round #2 and last) about the questionnaire http://www.henleymc.ac.uk/quest/40803.htm

- Valuate the scales

- **Does each scale measure what it is supposed to measure?**

- **Would kill/add/change any item? Pls. explain.**

Please, e-mail your feedback to me latest by xxxxxxxx. Let me know if that is impossible. Once again thank you very much for your help.

Best regards,

Raul M. Abril
HMC Research Associate
ra.dba15@henleymc.ac.uk

These are the scales (#1 to #17):

Scale #1 (Items #10 to #15): Customer Relationship Manager's perception on the extent to which the data warehouse customer relationship data is a source of customer linking information.

Scale #2 (Items #16 to #17): Customer Relationship Manager's perception on the extent to which the data warehouse customer relationship data is a source of competitor sensing information.

Scale #3 (Items #18 to #21): Customer Relationship Manager's perception on the extent to which the data warehouse customer relationship data is a source of applicable and helpful information to be used for understanding customer relationship problems.

Scale #4 (Items #22 to #26): Customer Relationship Manager's perception on the extent to which the data warehouse customer relationship data is integrated.

Scale #5 (Items #27, #29, #36): Customer Relationship Manager's perception on the extent to which the data warehouse customer relationship data favours more than one interpretation of customer relationship problems (equivocality).

Scale #6 (Items #30, #32 to #34, and #37): Customer Relationship Manager's perception on the extent to which the data warehouse customer relationship data favours to impart meaning when you have to understand customer relationship problems (richness).

Scale #7 (Items #28, #31, #35, #38, and #39): Customer Relationship Manager's perception on the extent to which the data warehouse customer relationship data is regarded as true and credible evidence for the understanding of customer relationship problems (Trustworthiness).

Scale #8 (Items #40 to #43): Customer Relationship Manager's self-belief about the extent that knowledge on customer relationship problems has improved due to the data warehouse customer relationship data.

Scale #9 (Items #44 to #47): Customer Relationship Manager's self-belief about the extent that one's skills understanding customer relationship problems have improved due to the data warehouse customer relationship data.

Scale #10 (Items #48 to #49): Customer Relationship Manager's self-belief about the extent that (i) the level of comprehensiveness -i.e. number of factors- and (ii) connectedness -i.e. links among the factors- of the problem statements about customer relationship problems have improved due to data warehouse customer relationship data.

Scale #11 (Items #50 to #61): The strength in the Customer Relationship Manager's self-beliefs in his/her capabilities to execute given types of performances understanding prospective customer relationship problems.

Scale #12 (Items #62(i), #63(i), and #64(i)): Customer Relationship Manager's perceptions on his/her amount of effort scanning the data warehouse. (Scanning search is the proactive behavior people exhibit when they browse through information without a particular problem to solve).

Scale #13 (Items #62(ii), #63(ii), and #64(ii)): Customer Relationship Manager's perceptions on his/her amount of effort in focused search on the data warehouse understanding customer relationship problems. Focused search is the reactive behavior people exhibit when they are looking for information specific to a problem to be addressed or question to be answered.

Scale #14 (Items #65(i), #66(i), #67(i) and #68(i)): Customer Relationship Manager's perceptions on his/her amount of effort using representativeness heuristic searching the data warehouse understanding customer relationship problems. Representativenes heuristic refers (i) To assess the probability of a situation as a representative of a

category (e.g. a 'price-lowering by a competitor' situation can have common information with an 'attempt action to gain market-share' pattern), or (ii) to make generalizations based on new information about a sample -i.e. the sample is representative of a large population- (e.g. to assess the national market success of a new product line based on the data likehood ratio of a test market). This heuristic implies a search for a satisficing fit bewteen information about a situation and information about a category.

Scale #15 (Items #65(ii), #66(ii), #67(ii) and #68(ii)): Customer Relationship Manager's perceptions on his/her amount of effort using availability heuristic searching the data warehouse understanding customer relationship problems. Availability heuristic refers to assess the probability of a situation as a function of prior situations (e.g. a marketer considering a series of occurences of actual costs incurred in past editions of a campaign when estimating the cost for a new edition of such campaign). This heuristic implies a search for recent, salient, easily accesible information about relevant precedents.

Scale #16 (Items #65(iii), #66(iii), #67(iii) and #68(iii)): Customer Relationship Manager's perceptions on his/her amount of effort using anchoring and adjustment heuristic searching the data warehouse understanding customer relationship problems. Anchoring and adjustment heuristic refers essentially to the trial and error method (e.g. a marketer setting the price of a product starting with a baseline price and making a number of impact analysis in several of the cost components). This heuristic implies a recursive process and each step involves a search for additional information and an adjustment of the previous assessment.

Scale #17 (Items #65(iv), #66(iv), #67(iv) and #68(iv)): Customer Relationship Manager's perceptions on his/her amount of effort using positivy heuristic searching the data warehouse understanding customer relationship problems. Positivy heuristic refers to confirm the probability of a situation (e.g. a marketer looking for issues in accounts that were predicted to have a high risk). This heuristic implies a search for information that is fundamentally consistent with existing beliefs, theories and cognition.

QUESTIONNAIRE.v2

Data Warehousing and Customer Relationship Management

First, thank you very much for your help. By answering this questionnaire, you can help marketers to more effectively execute CRM tasks. The questions focus on the quality your firm's data warehouse (i.e., integrated collection of customer data from various sources) and ways to enhance the benefits of CRM activities. Hopefully, you can easily an questions based on your experience in dealing with customer relationship issues. Based on our pretest, it should take no more than 20 minutes to complete the survey.

There are no right or wrong answers. Your responses will be kept in complete confidence and only be reported in a composite form with those of other respondents (i.e. no refe will be made to specific individuals or names of organizations).

If you have any questions or concerns please contact me using the e-mail address listed below.

Best regards,

Raul M. Abril
Research Associate
Henley Management College, UK
ra.dba15@henleymc.ac.uk

N. B.: Raul is following a doctoral program at Henley Management College, UK. His research focuses on the relationships between managerial problem solving, information search beh customer relationship management processes supported by a data warehouse, and data quality. As a practitioner, he has been working in the area of decision support systems for eigh years.

DETAILS ABOUT YOU AND YOUR CRM WORK ENVIRONMENT

The first 9 questions relate to your experiences in a CRM work environment. Please respond using the scale to the right of each question.

	Less than 6 months	Between 6 months and 1 year	Between 1 and 3 years	Between 3 and 5 years	More tha 5 years
1. **How long** have you been involved in customer relationship management tasks?	○	○	○	○	○
2. How long have you been using your **company's data warehouse customer relationship data** to help you understand issues?	○	○	○	○	○
3. **How long** has your data warehouse been supporting CRM functions (any of them)?	○	○	○	○	○

	No time at all									All o
	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
4. In a **typical week** how much time do you spend **understanding CRM issues** (e.g. challenges, problems)	○	○	○	○	○	○	○	○	○	○
5. In a **typical week** how much time do you spend **searching/requesting customer information** from your firm's data warehouse?	○	○	○	○	○	○	○	○	○	○

Sales Support = sales analysis, sales history, contact management, lead tracking, etc.

Marketing Support = customer analysis, channel analysis, channel management, campaign management, etc.

Customer Service Support = Call management, problem tracking, customer history, etc.

	Sales support	Marketing support	Customer service support	Other CRM functions
6. Which CRM functions are supported by your data warehouse? Check all that apply.	☐	☐	☐	☐
7. What is the **primary** CRM function of **your personal** job?.	○	○	○	○

To what extent do you agree or disagree with the following statements.	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Stron agre
8. My firm's data warehouse is an **important source** of customer relationship data for my personal job.	○	○	○	○	○
9. My firm's CRM activities target **individuals** as customers (i.e., not just other businesses or channel organizations)	○	○	○	○	○

PART I: YOUR FIRM'S DATA WAREHOUSE

Below there are 8 statements about the **actual** contents of your data warehouse in terms of customer relationship types of data. For each, please indicate the extent to which feel it reflects the situation in your organization. Click on a number on the scale to identify the degree to which you agree or disagree with the statement. **To what extent do y agree or disagree on the following statements about your firm's data warehouse?**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Stro ag
10. **It includes** customer's **motivational data** (e.g.needs)	○	○	○	○	○	○	C
11. **It includes** customer's **attitudinal data** (e.g. trust, satisfaction)	○	○	○	○	○	○	C
12. **It does not include relationship measures** (e.g. attrition risk, life-time value)	○	○	○	○	○	○	C
13. **It stores** customer **relationship events** (e.g., transactions, contacts, complaints)	○	○	○	○	○	○	C
14. **There is** customer's **geodemographic** data (e.g., age, gender)	○	○	○	○	○	○	C
15. **It does not include** customer's **behavioral** data (e.g., dominant purchase tendency, dominant channel usage habits)	○	○	○	○	○	○	C
16. **Competitors'** data **is not included** (e.g., products, campaigns, channels)	○	○	○	○	○	○	C
17. It is a source of data about **opportunities for competitive advantage**	○	○	○	○	○	○	C

Remember, we focus on *understanding* (i.e., not in *solving*). Below are two questions about the usefulness of your data warehouse's customer relationship data for understand customer relationship problems. **Please, give your opinion on ....**

	Not at all	A little extent	Some extent	Large extent	Very relevant
18. How **relevant** for understanding customer relationship problems is the customer relationship data in your firm's data warehouse?	○	○	○	○	○
	Never	**Rarely**	**Sometimes**	**Frequently**	**Most of the tin**
19. **How often is** the data you get from your data warehouse exactly what you require to understand customer relationship problems?	○	○	○	○	○

Below there are 20 statements about some characteristics of the customer relationship data in your data warehouse. Please respond as it describes the situation in your firm. a number on the scale to identify the degree to which you agree or disagree with the statement. **Your firm's data warehouse customer relationship data . . .**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Str ag
20. **enables** you **to find out the relevant factors** (i.e., components) of customer relationship problems.	○	○	○	○	○	○	(
21. does **not help** you to find out the **relevant relationships between the factors** (i.e., components) in customer relationship problems.	○	○	○	○	○	○	(
22. **enables** you to **easily identify the number of customers** your organization has.	○	○	○	○	○	○	(
23. **integrates** historic and operational data **at the customer level.**	○	○	○	○	○	○	(
24. **provides little help** in developing **a consistent 360° view** of each customer.	○	○	○	○	○	○	(
25. **partially** integrates the **relevant** customer relationship data from systems that serve different functional areas (e.g. sales, customer service).	○	○	○	○	○	○	(
26. **effectively integrates** customer relationship data from **a variety of data sources.**	○	○	○	○	○	○	(
27. **can be interpreted** in several ways that **lead** to different but acceptable customer relationship **problem formulations** .	○	○	○	○	○	○	(
28. includes incorrect **data.**	○	○	○	○	○	○	(
29. **means** different things to different people.	○	○	○	○	○	○	(
30. can **easily be summarized** at different levels.	○	○	○	○	○	○	(
31. has **missing data.**	○	○	○	○	○	○	(
32. **is organized** in a **meaningful** way.	○	○	○	○	○	○	(

Your firm's data warehouse customer relationship data	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Str a
33. does **not include** useful **information** for understanding customer relationship problems.	○	○	○	○	○	○	(
34. provides **new/unanticipated insights** for understanding customer relationship problems.	○	○	○	○	○	○	(
35. is a **trusted source** for understanding customer relationship problems.	○	○	○	○	○	○	(
36. can support **more than one plausible formulation** for the customer relationship problems you face.	○	○	○	○	○	○	(
37. challenges **a-priori characterizations** of customer relationship problems.	○	○	○	○	○	○	(
38. is **accurate.**	○	○	○	○	○	○	(
39. requires **double-checks** with other sources.	○	○	○	○	○	○	(

PART II: UNDERSTANDING CUSTOMER RELATIONSHIP PROBLEMS

Below there are 10 statements that relate to your understanding of customer relationship issues. Please indicate the extent to which each statement reflects the situation in yc personal case.

Think about the customer relationship problems you have faced in the last 6 months. **To what extent do you agree or disagree that the use of data warehouse custome relationship data has personally impacted your understanding of customer relationship problems?**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Str a
40. I have a **clearer sense** of customer relationship problems (i.e., better focus).	○	○	○	○	○	○	(
41. I reach **sharper final interpretations** of customer relationship problems.	○	○	○	○	○	○	(
42. I have **less understanding** of important customer relationship trends.	○	○	○	○	○	○	(

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Str a
43. I have **better insights** into the customer relationship threats facing my firm.	○	○	○	○	○	○	(
44. I look **for more information** in understanding customer relationship problems.	○	○	○	○	○	○	(
45. I analyze **more relevant information** before formulating a plausible customer relationship problem statement.	○	○	○	○	○	○	(
46. Quantitative analytic techniques are **more important to me** in understanding a customer relationship problem.	○	○	○	○	○	○	(
47. I am **less effective focusing** my attention on crucial information.	○	○	○	○	○	○	(
48. My final characterizations of customer relationship problems have **more factors** (i.e., problem components)	○	○	○	○	○	○	(
49. I do **not** reach final characterizations of customer relationship problems with **more interconnections among the factors.**	○	○	○	○	○	○	(

Some situations can make it harder or easier to define a plausible customer relationship problem. Below there are 12 statements relating to your degree of confidence in a num tasks. **Please rate your degree of confidence in your ability to complete the following tasks.**

	No chance					A 50-50 chance					Comp cer
	0	1	2	3	4	5	6	7	8	9	1
50. I can recognize **known factors** contributing to a **familiar** customer relationship problem.	○	○	○	○	○	○	○	○	○	○	(
51. I can formulate a plausible customer relationship problem statement when **I have access to the needed data.**	○	○	○	○	○	○	○	○	○	○	(
52. I can assess the **answerability of familiar questions** .	○	○	○	○	○	○	○	○	○	○	(
53. I can confirm conclusions for **familiar problems** looking at customer relationship data.	○	○	○	○	○	○	○	○	○	○	(

	No chance					A 50-50 chance					Comple certa
	0	1	2	3	4	5	6	7	8	9	10
54. I can identify **new factors** contributing to a customer relationship problem.	○	○	○	○	○	○	○	○	○	○	○
55. I can reach a plausible customer relationship problem statement under **moderate time pressure.**	○	○	○	○	○	○	○	○	○	○	○
56. I can identify **new questions** whose answers might give new insights.	○	○	○	○	○	○	○	○	○	○	○
57. I can explain cause-effect relationships in a problem when I have **limited information.**	○	○	○	○	○	○	○	○	○	○	○
58. I can understand **radically new types** of customer relationship problems.	○	○	○	○	○	○	○	○	○	○	○
59. I can develop a plausible statement about a **high stakes** (e.g., risk of losing a market segment) customer relationship problem.	○	○	○	○	○	○	○	○	○	○	○
60. I can explore **different patterns or trends** in customer relationship data.	○	○	○	○	○	○	○	○	○	○	○
61. I can analyze customer relationship data in a situation where there are **competing goals/objectives** about the desired customer relationship status (e.g., keep or terminate customer relationship).	○	○	○	○	○	○	○	○	○	○	○

PART III: INFORMATION SEARCH/REQUEST BEHAVIOR

62. Think of the situations in the last 6 months where you have been engaged in **searching/requesting customer relationship information** in your firm's data warehouse. Pl allocate 100 points across the following two situations to indicate **the relative frequency that you encounter these situations.** For example, if a 25 percent of your **searches/requests** are for information on "broad trend information", please, place a 25 beside such situation. The total should be 100.

(i) Searches/requests for information on **broad trend information** [ ]

(ii) Searches/requests for information to help you to understand **specific customer problems** or for **specific account information** [ ]

Total 100

63. Again, think of the situations in the last 6 months where you have been engaged in **searching/requesting customer relationship information** in your firm's data wareho
Please allocate 100 points across the following two situations to indicate **the relative frequency that you encounter these situations** . For example, if 80 percent of your **searches/requests** are for information in situations where you "react to a problem", please, place an 80 beside such situation. The total should be 100.

(i) Searches/requests for information on **unknown problems/key issues** [ ]

(ii) Searches/requests for information to help you **react** to a problem [ ]

Total 100

Listed below are 2 statements about the percentage of your time you search/request information in either an exploratory way or a focused way. For these questions, please us
following definitions.

Scanning means proactive browsing through information **without a particular problem** to solve (e.g., routine performance monitoring of a campaign).

Focused search/request means reactive looking for information **specific to a problem** to be addressed or question to be answered (e.g., information regarding a dissatisfied customer).

64. Please allocate 100 points across the following two searching situations to indicate **the ratio of your time spent** in a typical week in such situations. For example, if 80 p
your **time** searching for information it is spent in "scanning", please, place an 80 beside such a situation. The total should be 100.

(i) on data warehouse **scanning** . [ ] %

(ii) on data warehouse **focused** search/request. [ ] %

Total 100%

65. Think of the 'rules of thumb' that you have applied in the last 6 months in order to simplify your search/request for information in your firm's data warehouse when you need t understand a customer relationship problem. Please, allocate 100 points across the following five information search/request strategies to indicate **their relative frequency** . F example, if in 40 percent of your **searches/requests** you use "The best problem template at hand", please, place a 40 beside such strategy. The total should be 100.

(i) "The best 'problem template' at hand": Looking for **common information** between the situation that you have to understand and one or several **'problem template(s)'** that you have at hand	
ii) "The last occurrence" : Looking for information about **recent relevant precedent** situations	
(iii) "The trial and error": Adjusting your next question with the information that you got in a previous question	
(iv) "The confirmatory": Looking for information that **confirms** your belief about the problem	
(v) Other information search/request strategies	
Total	100

66. Think of the 'rules of thumb' that you have applied in the last 6 months in order to simplify your search/request for information in your firm's data warehouse when you need t understand a customer relationship problem. Please, allocate 100 points across the following five information search/request strategies to indicate **their relative frequency.** F example, if in a 60 percent of your **searches/requests** you use "The diagnostic", please, place a 60 beside such strategy. The total should be 100.

(i) "The test" : Looking for information in order to identify an acceptable **sample/target** in order to be able to generalize to a larger scale	
(ii) "The 'canned' queries" : Using **predefined** questions	
(iii) "The divide & conquer" : **Adding** successive selection **criteria** to your initial question	
(iv) "The diagnostic" : Looking for information that **confirms** known symptoms.	
(v) Other information search/request strategies	
Total	100

67. Please allocate 100 points across the following five searching "rules of thumb" to indicate **the ratio of your time spent** in a typical week using such strategies. For examp 80 percent of your **time** searching for information is spent using "the trial and error" strategy, please, place a 80 beside such strategy. The total should be 100.

(i) Using "The best 'problem template' at hand" strategy		%
(ii) Using "The last occurrence" strategy		%
(iii) Using "The trial and error" strategy		%
(iv) Using "The confirmatory" strategy		%
(v) Using other information search/request strategies		%
Total	100%	

68. Again, please allocate 100 points across the following four searching "rules of thumb" to indicate **the ratio of your time spent** in a typical week using such strategies. For example, if 80 percent of your **time** searching for information it is spent using "the 'canned' queries" strategy, please, place a 80 beside such strategy. The total should be 100

(i) Using "The test" search/request strategy		%
(ii) Using "The 'canned' queries" search/request strategy		%
(iii) Using "The divide and conquer" search/request strategy		%
(iv) Using "The diagnostic" search/request strategy		%
(v) Using other information search/request strategies		%
Total	100%	

RESPONDENT'S PERSONAL DATA

The following information will only be used to provide an overall picture of the total sample. All individual or company specific data will be retained in the strictest confidence. **All marked with an asterisk are optional and respondents will only be followed up personally when they have indicated a willingness to do so.**

Name

Position Held

Function

- ○ Bus. Development
- ○ Customer Service (e.g. Call Center)
- ○ E-Business
- ○ Executive Office
- ○ Fraud Management
- ○ Marketing - General
- ○ Marketing - CRM (e.g. Research, Channel, Campaign, Segment)
- ○ Merchandising
- ○ Sales (e.g. Region, Account)
- ○ Strategic Planning
- ○ Other (please specify)

Level*

- ○ Board Member (Chairman, VC)
- ○ Business Head (CEO, President)
- ○ Bus Function Leader (CIO, CMO)
- ○ Sr. Dept Head (Director, AVP, VP)
- ○ Other Dept/Function Head (Manager)
- ○ Project/Team Leader (Project Mgr)
- ○ Analyst
- ○ Individual Contributor (e.g. Representative)
- ○ Other (Please specify)

Organization / Company*

What is your data warehouse' s database? *

- ○ DB2
- ○ Teradata
- ○ Oracle
- ○ Other (Please specify) [ ]
- ○ SQL-Server

What is the primary business of your company? *

- ○ Manufacturing
- ○ Pharmaceuticals
- ○ Retail
- ○ Chemical
- ○ Financial/Insurance
- ○ Computer & IT
- ○ Telecommunications
- ○ Other (Please specify) [ ]

Would you like to receive a summary of the research findings?

- ○ Yes
- ○ No

If "YES"

Telephone contact [ ]

e-mail address [ ]

THANK YOU, FOR TAKING THE TIME TO RESPOND TO THIS SURVEY.
Please scroll down to the bottom of the page and click on the submit button.

Submit Reset

Created with HTML +Forms

APPENDIX D: PILOT RESEARCH STUDY.

QUESTIONNAIRE.v3

N: No action (i.e. scale unchanged)
C#: I changed # items
A#: I added # items
K# I dropped # items

+ Alpha in survey Improved
= Alpha in survey at the same level
- Alpha in survey dropped

	SCALE	Pilot				Sample	
		Alpha	# Items	Changed /Added /Killed	Decision's Impact on sample	Alpha	# Items
CUST	Customer Insights	0.74	6	N	-	0.47	6
COMP	Competitor Insights	-0.13	2	C1 A1	+	0.71	3
ENAC	Data Instrumental utility (customer relationship problem enactment)	0.84	4	N	=	0.83	4
INT	Data Integration	-0.08	5	C3 K2	+	0.58	3
EQUI	Data Equivocality	0.73	3	N	-	0.51	3
RICH	Data Richness	0.77	5	C1	=	0.78	5
TRUST	Data Trustworthiness	0.60	5	C2 K2	=	0.55	3

COMP: I removed the code reverse of item_7.v3. I introduced a third item from the original Competitor Orientation scale that I did not consider the first time.
INT: I reworded 3 items (removing the two code reverse wordings), and dropped two items
RICH: I reworded item item_21.v3
TRUST: I reworded items item_19.v3 and item_22.v3 (removing their code reverse), and dropped items item_29.v3 and item_30.v3.

	SCALE	Pilot				Sample	
		Alpha	# Items	Changed /Added /Killed	Decision's Impact on sample	Alpha	# Items
SCAN/FOCUS	Intensity of data warehouse scanning	0.28	3	C1 K1	+	0.66	2
REP	Intensity of data warehouse representativeness	0.92	4	N	-	0.66	4
AVBLE	Intensity of data warehouse availability search heuristic	0.61	4	C2	=	0.67	4
AA	Intensity of data warehouse anchoring and adjustment search	0.93	4	N	-	0.78	4
POSI	Intensity of data warehouse positivy search heuristic	0.90	4	N	-	0.64	4

SCAN: I dropped item_53.v3i and reworded item_55.v3i
FOCUS: I dropped item_53.v3ii and reworded item_55.v3ii
AVBLE: I reworded 2 items item_56ii.v3 and item_58ii.v3

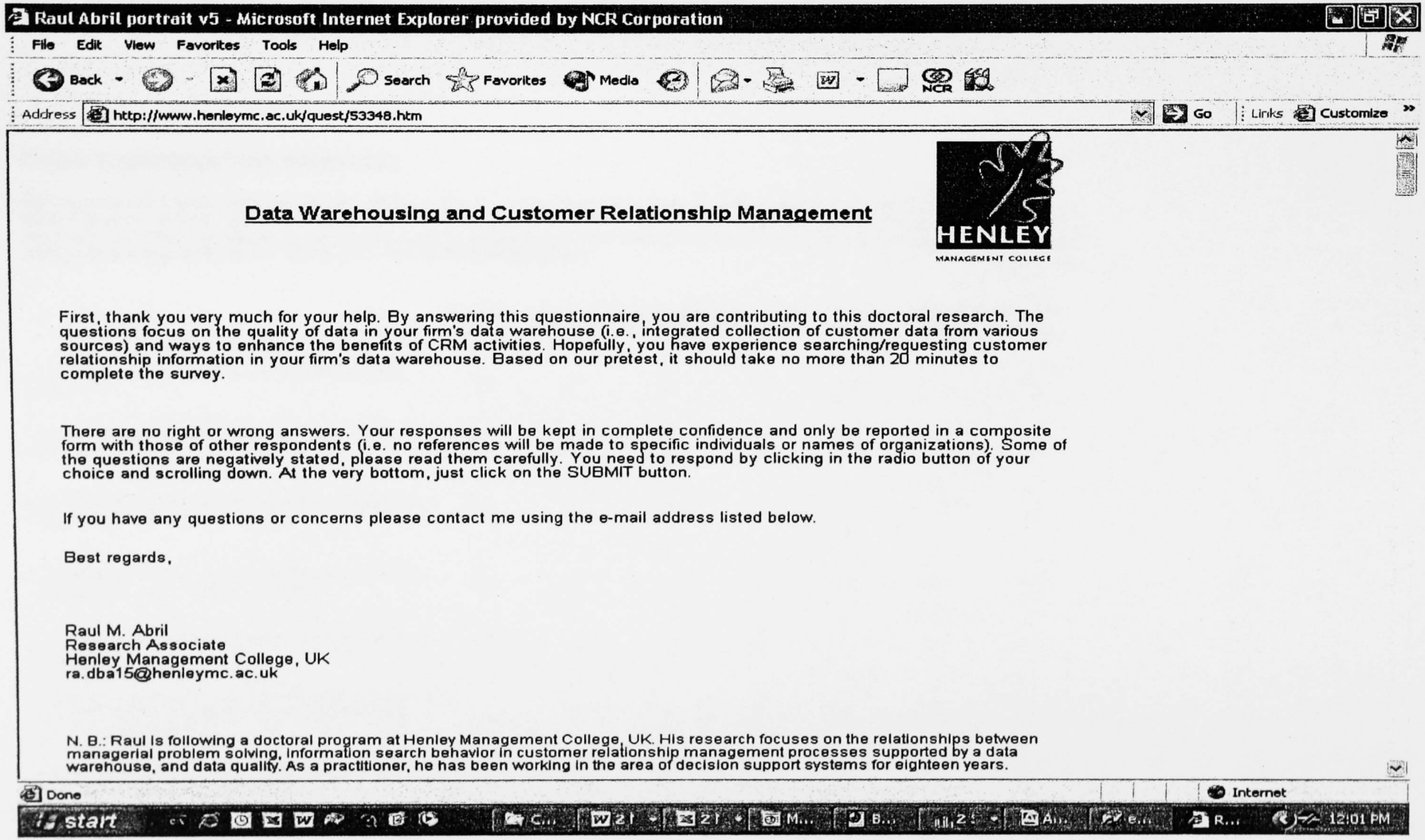
N: No action (i.e. scale unchanged)
C#: I changed # items
A#: I added # items
K# I dropped # items

+ Alpha in survey Improved
= Alpha in survey at the same level
- Alpha in survey dropped

	SCALE	Pilot				Sample	
		Alpha	# Items	Changed /Added /Killed	Decision's Impact on sample	Alpha	# Items
DKNOWE	Customer Relationship Problem Declarative Knowledge Enhancement	0.86	4	K1	+	0.95	3
PKNOWE	Customer Relationship Problem Enactment Procedural Knowledge Enhancement	0.49	4	C4	+	0.59	4
ICPLXE	Customer Relationship Problem Integrative Complexity Enhancement	0.51	2	C2	+	0.82	2
LSELFE	Low Customer Relationship Problem Enacting Self-Efficacy	0.62	4	C4	+	0.82	4
MSELFE	Medium Customer Relationship Problem Enacting Self-Efficacy	0.83	4	C1	=	0.84	4
HSELFE	High Customer Relationship Problem Enacting Self-Efficacy	0.84	4	N	=	0.80	4
SELFE	Customer Relationship Problem Enacting Self-Efficacy	0.90	12		=	0.92	12

DKNOWE: I dropped item_.33v3 (code reverse)
PKNOWE: I reworded items item_35.v3, item_36.v3 and item_37.v3 and dropped the reverse code in item_38.v3.
ICPLXE: I reworded both items and removed the code reverse in item_40.v3
LSELFE: I reworded the 4 items
MSELFE: I reworded 1 item_48.v3

QUESTIONNAIRE.v3



Data Warehousing and Customer Relationship Management

HENLEY
MANAGEMENT COLLEGE

First, thank you very much for your help. By answering this questionnaire, you are contributing to this doctoral research. The questions focus on the quality of data in your firm's data warehouse (i.e., integrated collection of customer data from various sources) and ways to enhance the benefits of CRM activities. Hopefully, you have experience searching/requesting customer relationship information in your firm's data warehouse. Based on our pretest, it should take no more than 20 minutes to complete the survey.

There are no right or wrong answers. Your responses will be kept in complete confidence and only be reported in a composite form with those of other respondents (i.e. no references will be made to specific individuals or names of organizations). Some of the questions are negatively stated, please read them carefully. You need to respond by clicking in the radio button of your choice and scrolling down. At the very bottom, just click on the SUBMIT button.

If you have any questions or concerns please contact me using the e-mail address listed below.

Best regards,

Raul M. Abril
Research Associate
Henley Management College, UK
ra.dba15@henleymc.ac.uk

N. B.: Raul is following a doctoral program at Henley Management College, UK. His research focuses on the relationships between managerial problem solving, information search behavior in customer relationship management processes supported by a data warehouse, and data quality. As a practitioner, he has been working in the area of decision support systems for eighteen years.

N. B.: Raul is following a doctoral program at Henley Management College, UK. His research focuses on the relationships between managerial problem solving, information search behavior in customer relationship management processes supported by a data warehouse, and data quality. As a practitioner, he has been working in the area of decision support systems for eighteen years.

PART I: YOUR FIRM'S DATA WAREHOUSE

Below there are 8 statements about the **actual** contents of your data warehouse in terms of customer relationship types of data. For each, please indicate the extent to which you feel it reflects the situation in your organization. Click on a number on the scale to identify the degree to which you agree or disagree with the statement. **To what extent do you agree or disagree on the following statements about your firm's data warehouse?**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
1. **It includes** customer's **motivational data** (e.g.needs)	○	○	○	○	○	○	○
2. **It includes** customer's **attitudinal data** (e.g. trust, satisfaction)	○	○	○	○	○	○	○
3. **It does not include relationship measures** (e.g. attrition risk, life-time value)	○	○	○	○	○	○	○
4. **It stores** customer **relationship events** (e.g., transactions, contacts, complaints)	○	○	○	○	○	○	○
5. **There is** customer's **demographic** data (e.g., age, gender)	○	○	○	○	○	○	○
6 **It does not include** customer's **behavioral** data (e.g., dominant purchase tendency, dominant channel usage habits)	○	○	○	○	○	○	○
7. **Competitors'** data **is not included** (e.g., products, campaigns, channels)	○	○	○	○	○	○	○
8. It is a source of data about **opportunities for competitive advantage**	○	○	○	○	○	○	○

8. It is a source of data about **opportunities for competitive advantage**	○	○	○	○	○	○	○

Remember, we focus on *understanding* (i.e., not in *solving*). Below are two questions about the usefulness of your data warehouse's customer relationship data for <u>understanding</u> customer relationship problems. **Please, give your opinion on ....**

	Not at all	A little extent	Some extent	Large extent	Extremely
9. How **relevant** for understanding customer relationship problems is the customer relationship data in your firm's data warehouse?	○	○	○	○	○

	Never	Rarely	Sometimes	Most of the time	All the time
10. **How often is** the data you get from your data warehouse exactly what you require to understand customer relationship problems?	○	○	○	○	○

Below there are 20 statements about some characteristics of the customer relationship data in your data warehouse. Please respond as it describes the situation in your firm. Click on a number on the scale to identify the degree to which you agree or disagree with the statement. **Your firm's data warehouse customer relationship data . . .**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
11. **enables** you **to find out the relevant factors** (i.e., components) of customer relationship problems.	○	○	○	○	○	○	○
12. does **not help** you to find out the **relevant relationships between the factors** (i.e., components) in customer relationship problems.	○	○	○	○	○	○	○
13. **enables** you to **easily identify the number of customers** your organization has.	○	○	○	○	○	○	○
14. **integrates** historic and operational timeliness data **at the customer level.**	○	○	○	○	○	○	○
15 **provides little help** in developing **a consistent 360° view** of each customer.	○	○	○	○	○	○	○
16. **partially** integrates the **relevant** customer relationship data from systems that serve different functional areas (e.g. sales, customer service).	○	○	○	○	○	○	○

15 **provides little help** in developing **a consistent 360° view** of each customer.	○	○	○	○	○	○	○
16. **partially** integrates the **relevant** customer relationship data from systems that serve different functional areas (e.g. sales, customer service).	○	○	○	○	○	○	○
17 **effectively integrates** customer relationship data from **a variety of data sources.**	○	○	○	○	○	○	○
18. **can be interpreted** in several ways that **lead** to different but acceptable customer relationship **problem formulations** .	○	○	○	○	○	○	○
19. includes incorrect **data.**	○	○	○	○	○	○	○
20. **means** different things to different people.	○	○	○	○	○	○	○
	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
21. can **easily be summarized** at different levels.	○	○	○	○	○	○	○
22. has **missing data.**	○	○	○	○	○	○	○
23. **is organized** in a **meaningful** way.	○	○	○	○	○	○	○
24. does **not include** useful **information** for understanding customer relationship problems.	○	○	○	○	○	○	○
25. provides **new/unanticipated insights** for understanding customer relationship problems.	○	○	○	○	○	○	○
26. is a **trusted source** for understanding customer relationship problems.	○	○	○	○	○	○	○
27. can support **more than one plausible formulation** for the customer relationship problems you face.	○	○	○	○	○	○	○
28. challenges **a-priori characterizations** of customer relationship problems.	○	○	○	○	○	○	○
29. is **accurate.**	○	○	○	○	○	○	○
30. requires **double-checks** with other sources.	○	○	○	○	○	○	○

30. requires **double-checks** with other sources.	○	○	○	○	○	○	○

PART II: UNDERSTANDING CUSTOMER RELATIONSHIP PROBLEMS

Below there are 10 statements that relate to your understanding of customer relationship issues. Please indicate the extent to which each statement reflects the situation in your personal case.

Think about the customer relationship problems you have faced in the last 6 months. **To what extent do you agree or disagree that the use of data warehouse customer relationship data has personally impacted your understanding of customer relationship problems?**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
31. I have a **clearer sense** of customer relationship problems (i.e., better focus).	○	○	○	○	○	○	○
32. I reach **sharper final interpretations** of customer relationship problems.	○	○	○	○	○	○	○
33. I have **less understanding** of important customer relationship trends.	○	○	○	○	○	○	○
34. I have **better insights** into the customer relationship threats facing my firm.	○	○	○	○	○	○	○
35. I search **for more information** in understanding customer relationship problems.	○	○	○	○	○	○	○
36. I analyze **more relevant information** before formulating a plausible customer relationship problem statement.	○	○	○	○	○	○	○
37. Quantitative analytic techniques are **more important to me** in understanding a customer relationship problem.	○	○	○	○	○	○	○
38. I am **less effective focusing** my attention on crucial information.	○	○	○	○	○	○	○
39. My final characterizations of customer relationship problems have **more factors** (i.e., problem components)	○	○	○	○	○	○	○
40. I do **not** reach final characterizations of customer relationship problems with **more interconnections among the factors.**	○	○	○	○	○	○	○

40. I do **not** reach final characterizations of customer relationship problems with **more interconnections among the factors.** ○ ○ ○ ○ ○ ○ ○

Some situations can make it harder or easier to define a plausible customer relationship problem. Below there are 12 statements relating to your degree of confidence in a number of tasks. **Please rate your degree of confidence in your ability to complete the following tasks.**

	No chance					A 50-50 chance					Completely certain
	0	1	2	3	4	5	6	7	8	9	10
41. I can recognize **known factors** contributing to a **familiar** customer relationship problem.	○	○	○	○	○	○	○	○	○	○	○
42. I can formulate a plausible customer relationship problem statement when **I have access to the needed data.**	○	○	○	○	○	○	○	○	○	○	○
43. I can assess the **answerability of familiar questions** .	○	○	○	○	○	○	○	○	○	○	○
44. I can confirm conclusions for **familiar problems** looking at customer relationship data.	○	○	○	○	○	○	○	○	○	○	○
45. I can identify **new factors** contributing to a customer relationship problem.	○	○	○	○	○	○	○	○	○	○	○

	No chance					A 50-50 chance					Completely certain
	0	1	2	3	4	5	6	7	8	9	10
46. I can reach a plausible customer relationship problem statement under **moderate time pressure.**	○	○	○	○	○	○	○	○	○	○	○
47. I can identify **new questions** whose answers might give new insights.	○	○	○	○	○	○	○	○	○	○	○
48. I can explain cause-effect relationships in a problem when I have **limited information.**	○	○	○	○	○	○	○	○	○	○	○
49. I can understand **radically new types** of customer relationship problems.	○	○	○	○	○	○	○	○	○	○	○
50. I can develop a plausible statement about a **high stakes** (e.g., risk of losing a market segment) customer relationship problem.	○	○	○	○	○	○	○	○	○	○	○

51. I can explore **different patterns or trends** in customer relationship data. ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○

52. I can analyze customer relationship data in a situation where there are **competing goals/objectives** about the desired customer relationship status (e.g., keep or terminate customer relationship). ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○

PART III: INFORMATION SEARCH/REQUEST BEHAVIOR

53. Think of the situations in the last 6 months where you have been engaged in **searching/requesting customer relationship information** in your firm's data warehouse. Please allocate 100 points across the following two situations to indicate **the relative frequency that you encounter these situations.** For example, if a 25 percent of your **searches/requests** are for information on "broad trend information", please, place a 25 beside such situation. The total should be 100.

(i) Searches/requests for information on **broad trend information** [ ]

(ii) Searches/requests for information to help you to understand **specific customer problems** or for **specific account information** [ ]

Total 100

54. Again, think of the situations in the last 6 months where you have been engaged in **searching/requesting customer relationship information** in your firm's data warehouse. Please allocate 100 points across the following two situations to indicate **the relative frequency that you encounter these situations** . For example, if 80 percent of your **searches/requests** are for information in situations where you "react to a problem", please, place an 80 beside such situation. The total should be 100.

(i) Searches/requests for information on **unknown problems/key issues** [ ]

(ii) Searches/requests for information to help you **react** to a problem [ ]

Total 100

Listed below are 2 statements about the percentage of your time you search/request information in either an exploratory way or a focused way. For these questions, please use the following definitions.

Scanning means proactive browsing through information **without a particular problem** to solve (e.g., routine performance monitoring of a campaign).

Focused search/request means reactive looking for information **specific to a problem** to be addressed or question to be answered (e.g., information regarding a dissatisfied customer).

55. Please allocate 100 points across the following two searching situations to indicate **the ratio of your time spent** in a typical week in such situations. For example, if 80 percent of your **time** searching for information it is spent in "scanning", please, place an 80 beside such a situation. The total should be 100.

(i) on data warehouse **scanning** . [ ] %

(ii) on data warehouse **focused** search/request. [ ] %

Total 100%

56. Think of the 'rules of thumb' that you have applied in the last 6 months in order to simplify your search/request for information in your firm's data warehouse when you need to understand a customer relationship problem. Please, allocate 100 points across the following five information search/request strategies to indicate **their relative frequency** . For example, if in 40 percent of your **searches/requests** you use "The best problem template at hand", please, place a 40 beside such strategy. The total should be 100.

(i) "The best 'problem template' at hand": Looking for **common information** between the situation that you have to understand and one or several **'problem template(s)'** that you have at hand [ ]

ii) "The last occurrence" : Looking for information about **recent relevant precedent** situations [ ]

(iii) "The trial and error": Adjusting your next question with the information that you got in a previous question [ ]

(iv) "The confirmatory": Looking for information that **confirms** your belief about the problem [ ]

(v) Other information search/request strategies [ ]

Total 100

57. Please allocate 100 points across the following five searching "rules of thumb" to indicate **the ratio of your time spent** in a typical week using such strategies. For example, if 80 percent of your **time** searching for information is spent using "the trial and error" strategy, please, place a 80 beside such strategy. The total should be 100.

(i) Using "The best 'problem template' at hand" strategy		%
(ii) Using "The last occurrence" strategy		%
(iii) Using "The trial and error" strategy		%
(iv) Using "The confirmatory" strategy		%
(v) Using other information search/request strategies		%
Total	100%	

58. Think of the 'rules of thumb' that you have applied in the last 6 months in order to simplify your search/request for information in your firm's data warehouse when you need to understand a customer relationship problem. Please, allocate 100 points across the following five information search/request strategies to indicate **their relative frequency.** For example, if in a 60 percent of your **searches/requests** you use "The diagnostic", please, place a 60 beside such strategy. The total should be 100.

(i) "The test" : Looking for information in order to identify an acceptable **sample/target** in order to be able to generalize to a larger scale	
(ii) "The 'canned' queries" : Using **predefined** questions	
(iii) "The divide & conquer" : **Adding** successive selection **criteria** to your initial question	
(iv) "The diagnostic" : Looking for information that **confirms** known symptoms.	
(v) Other information search/request strategies	
Total	100

Total 100

59. Again, please allocate 100 points across the following four searching "rules of thumb" to indicate **the ratio of your time spent** in a typical week using such strategies. For example, if 80 percent of your **time** searching for information it is spent using "the 'canned' queries" strategy, please, place a 80 beside such strategy. The total should be 100

(i) Using "The test" search/request strategy		%
(ii) Using "The 'canned' queries" search/request strategy		%
(iii) Using "The divide and conquer" search/request strategy		%
(iv) Using "The diagnostic" search/request strategy		%
(v) Using other information search/request strategies		%
Total	100%	

RESPONDENT'S PERSONAL DATA

The following information will only be used to provide an overall picture of the total sample. All data will be retained in the strictest confidence. **Respondents will only be followed up personally when they have indicated a willingness to do so.**

	Less than 6 months	Between 6 months and 1 year	Between 1 and 3 years	Between 3 and 5 years	More than 5 years
60. **How long** have you been involved in customer relationship management tasks?	○	○	○	○	○
61. How long have you been using your **company's data warehouse customer relationship data** to help you understand issues?	○	○	○	○	○
62. **How long** has your data warehouse been supporting CRM functions (any of them)?	○	○	○	○	○

	No time at all										All of my time
	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
63. In a **typical week** how much time do you spend **understanding CRM issues** (e.g. challenges, problems)	○	○	○	○	○	○	○	○	○	○	○
64. In a **typical week** how much time do you spend **searching/requesting customer information** from your firm's data warehouse?	○	○	○	○	○	○	○	○	○	○	○

Sales Support = sales analysis, sales history, contact management, lead tracking, etc.

Marketing Support = customer analysis, channel analysis, channel management, campaign management, etc.

Customer Service Support = Call management, problem tracking, customer history, etc.

	Sales support	Marketing support	Customer service support	Other CRM functions
65. Which CRM functions are supported by your data warehouse? Check all that apply.	☐	☐	☐	☐
66. What is the **primary** CRM function of **your personal** job?.	○	○	○	○

To what extent do you agree or disagree with the following statements.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
67. My firm's data warehouse is an **important source** of customer relationship data for my personal job.	○	○	○	○	○	○	○
68. My firm's CRM activities target **individuals** as customers (i.e., not just other businesses or channel organizations)	○	○	○	○	○	○	○

67. My firm's data warehouse is an **important source** of customer relationship data for my personal job.	○	○	○	○	○	○	○
68. My firm's CRM activities target **individuals** as customers (i.e., not just other businesses or channel organizations)	○	○	○	○	○	○	○

Name

Organization / Company

Would you like to receive a summary of the research findings? ○ Yes ○ No

If "YES"

Telephone contact

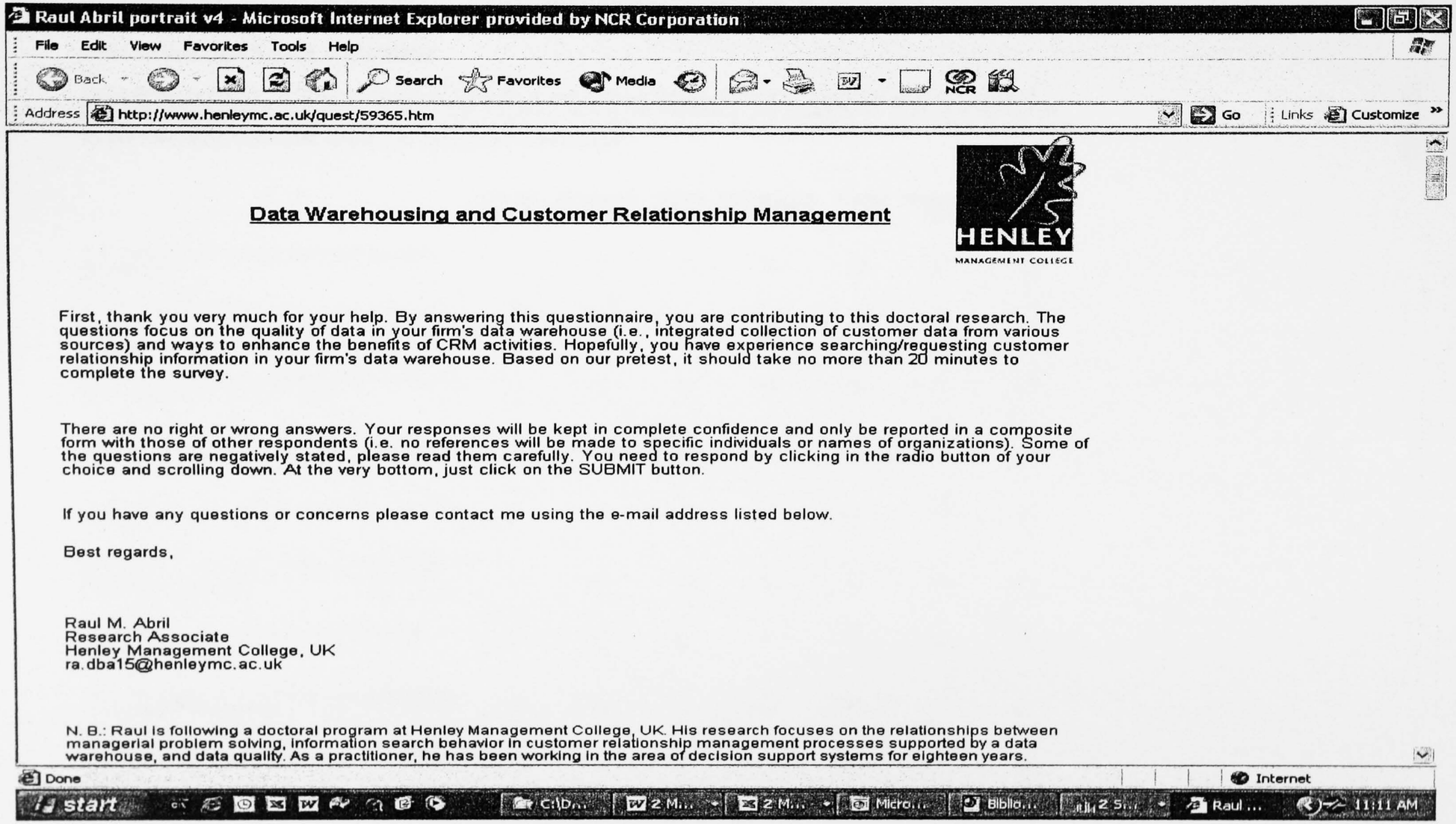
e-mail address

THANK YOU, FOR TAKING THE TIME TO RESPOND TO THIS SURVEY. Please scroll down to the bottom of the page and click on the submit button.

Submit Reset

Created with HTML+Forms

APPENDIX E: SURVEY QUESTIONNAIRE.v4



PART I: YOUR FIRM'S DATA WAREHOUSE

Below there are 9 statements about the **actual** contents of your data warehouse in terms of customer relationship types of data. For each, please indicate the extent to which you feel it reflects the situation in your organization. Click on a number on the scale to identify the degree to which you agree or disagree with the statement. **To what extent do you agree or disagree on the following statements about your firm's data warehouse?**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
1. **It includes** customer's **motivational data** (e.g.needs)	○	○	○	○	○	○	○
2. **It includes** customer's **attitudinal data** (e.g. trust, satisfaction)	○	○	○	○	○	○	○
3. **It does not include relationship measures** (e.g. attrition risk, life-time value)	○	○	○	○	○	○	○
4. **It stores** customer **relationship events** (e.g., transactions, contacts, complaints)	○	○	○	○	○	○	○
5. **There is** customer's **demographic** data (e.g., age, gender)	○	○	○	○	○	○	○
6 **It does not include** customer's **behavioral** data (e.g., dominant purchase tendency, dominant channel usage habits)	○	○	○	○	○	○	○
7. Data about **competitors** is included (e.g., products, campaigns, channels)	○	○	○	○	○	○	○
8. It is a source of data about **opportunities for competitive advantage**	○	○	○	○	○	○	○
9. Allows to **respond rapidly** to competitors' actions	○	○	○	○	○	○	○

Remember, we focus on *understanding* (i.e., not in *solving*). Below are two questions about the usefulness of your data warehouse's customer relationship data for understanding customer relationship problems. **Please, give your opinion on ....**

	Not at all	A little extent	Some extent	Large extent	Extremely
10. How **relevant** for understanding customer relationship problems is the customer relationship data in your firm's data warehouse?	○	○	○	○	○

	Never	Rarely	Sometimes	Most of the time	All the time
11. **How often is** the data you get from your data warehouse exactly what you require to understand customer relationship problems?	○	○	○	○	○

Below there are 16 statements about some characteristics of the customer relationship data in your data warehouse. Please respond as it describes the situation in your firm. Click on a number on the scale to identify the degree to which you agree or disagree with the statement. **Your firm's data warehouse customer relationship data . . .**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
12. **enables** you **to find out the relevant factors** (i.e., components) of customer relationship problems.	○	○	○	○	○	○	○
13. does **not help** you to find out the **relevant relationships between the factors** (i.e., components) in customer relationship problems.	○	○	○	○	○	○	○
14. allows you **to know the number of customers in your data warehouse** .	○	○	○	○	○	○	○
15. allows a **360° view** of a customer.	○	○	○	○	○	○	○
16. **integrates** customer data from **a variety of data sources.**	○	○	○	○	○	○	○
17. **can be interpreted** in several ways that **lead** to different but acceptable customer relationship **problem formulations** .	○	○	○	○	○	○	○
18. includes **accurate** data.	○	○	○	○	○	○	○

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
13. does **not help** you to find out the **relevant relationships between the factors** (i.e., components) in customer relationship problems.	○	○	○	○	○	○	○
14. allows you **to know the number of customers in your data warehouse** .	○	○	○	○	○	○	○
15. allows a **360° view** of a customer.	○	○	○	○	○	○	○
16. **integrates** customer data from **a variety of data sources.**	○	○	○	○	○	○	○
17. **can be interpreted** in several ways that **lead** to different but acceptable customer relationship **problem formulations** .	○	○	○	○	○	○	○
18. includes **accurate** data.	○	○	○	○	○	○	○
19. **means** different things to different people.	○	○	○	○	○	○	○
20. can **be summarized** at different levels.	○	○	○	○	○	○	○
21. has **missing data.**	○	○	○	○	○	○	○
22. **is organized** in a **meaningful** way.	○	○	○	○	○	○	○
23. does **not include** useful **information** for understanding customer relationship problems.	○	○	○	○	○	○	○
24. provides **new/unanticipated insights** for understanding customer relationship problems.	○	○	○	○	○	○	○
25. is a **trusted source** of customer data.	○	○	○	○	○	○	○
26. can support **more than one plausible formulation** for the customer relationship problems you face.	○	○	○	○	○	○	○
27. challenges **your initial characterizations** of customer relationship problems.	○	○	○	○	○	○	○

problems you face.

27. challenges **your initial characterizations** of customer relationship problems.	○	○	○	○	○	○	○

PART II: UNDERSTANDING CUSTOMER RELATIONSHIP PROBLEMS

Below there are 9 statements that relate to your understanding of customer relationship issues. Please indicate the extent to which each statement reflects the situation in your personal case.

Think about the customer relationship problems you have faced in the last 6 months . **To what extent do you agree or disagree that the use of data warehouse customer relationship data has personally impacted your understanding of customer relationship problems?**

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
28. I have a **clearer sense** of customer relationship problems (i.e., better focus).	○	○	○	○	○	○	○
29. I reach **sharper final interpretations** of customer relationship problems.	○	○	○	○	○	○	○
30. I have **better insights** into the customer relationship threats facing my firm.	○	○	○	○	○	○	○
31. I spend **more time** searching for key information.	○	○	○	○	○	○	○
32. I analyze **more** a customer relationship problem.	○	○	○	○	○	○	○
33. Quantitative analytic techniques are **more helpful** in understanding a customer relationship problem.	○	○	○	○	○	○	○
34. I am more effective **focusing on the key aspects** of customer relationship problems.	○	○	○	○	○	○	○
35. My descriptions of customer relationship problems include **more factors** .	○	○	○	○	○	○	○
36. My descriptions of customer relationship problems include **more interconnections among the factors.**	○	○	○	○	○	○	○

Some situations can make it harder or easier to define a plausible customer relationship problem. Below there are 12 statements relating to your degree of confidence in a number of tasks. **Please rate your degree of confidence in your ability to complete the following tasks.**

	No chance					A 50-50 chance					Completely certain
	0	1	2	3	4	5	6	7	8	9	10
37. I can recognize **common factors** contributing to a customer relationship problem.	○	○	○	○	○	○	○	○	○	○	○
38. I can formulate a plausible customer relationship problem statement when **I have the data that I need** .	○	○	○	○	○	○	○	○	○	○	○
39. I can determine if there is **available data to answer common questions** .	○	○	○	○	○	○	○	○	○	○	○
40. I can confirm my preliminary conclusions for **common problems** .	○	○	○	○	○	○	○	○	○	○	○
41. I can identify **new factors** contributing to a customer relationship problem.	○	○	○	○	○	○	○	○	○	○	○

	No chance					A 50-50 chance					Completely certain
	0	1	2	3	4	5	6	7	8	9	10
42. I can reach a plausible customer relationship problem statement under **moderate time pressure.**	○	○	○	○	○	○	○	○	○	○	○
43. I can identify **new questions** whose answers might give new insights.	○	○	○	○	○	○	○	○	○	○	○
44. I can explain cause-effect relationships in a problem even with **limited information.**	○	○	○	○	○	○	○	○	○	○	○
45. I can understand **radically new types** of customer relationship problems.	○	○	○	○	○	○	○	○	○	○	○
46. I can develop a plausible statement about a **high stakes** (e.g., risk of losing a market segment) customer relationship problem.	○	○	○	○	○	○	○	○	○	○	○

	No chance					A 50-50 chance					Completely certain
	0	1	2	3	4	5	6	7	8	9	10
42. I can reach a plausible customer relationship problem statement under **moderate time pressure.**	○	○	○	○	○	○	○	○	○	○	○
43. I can identify **new questions** whose answers might give new insights.	○	○	○	○	○	○	○	○	○	○	○
44. I can explain cause-effect relationships in a problem even with **limited information.**	○	○	○	○	○	○	○	○	○	○	○
45. I can understand **radically new types** of customer relationship problems.	○	○	○	○	○	○	○	○	○	○	○
46. I can develop a plausible statement about a **high stakes** (e.g., risk of losing a market segment) customer relationship problem.	○	○	○	○	○	○	○	○	○	○	○
47. I can explore **different patterns or trends** in customer relationship data.	○	○	○	○	○	○	○	○	○	○	○
48. I can analyze customer relationship data in a situation where there are **competing goals/objectives** about the desired customer relationship status (e.g., keep or terminate customer relationship).	○	○	○	○	○	○	○	○	○	○	○

PART III: INFORMATION SEARCH/REQUEST BEHAVIOR

49. Think of the situations in the last 6 months where you have been engaged in **searching/requesting customer relationship information** in your firm's data warehouse. Please allocate 100 points across the following two situations to indicate **the relative frequency that you encounter these situations** . For example, if 80 percent of your **searches/requests** are for information in situations where you "react to a problem", please, place an 80 beside such situation. The total should be 100. **How often do your searches/requests involve information...**

(i) on **unknown problems/key issues** ? ☐ %

(ii) to help you **react** to a problem? ☐ %

Total 100%

stakes (e.g., risk of losing a market segment) customer relationship problem.	○	○	○	○	○	○	○	○	○	○	○
47. I can explore **different patterns or trends** in customer relationship data.	○	○	○	○	○	○	○	○	○	○	○
48. I can analyze customer relationship data in a situation where there are **competing goals/objectives** about the desired customer relationship status (e.g., keep or terminate customer relationship).	○	○	○	○	○	○	○	○	○	○	○

PART III: INFORMATION SEARCH/REQUEST BEHAVIOR

49. Think of the situations in the last 6 months where you have been engaged in **searching/requesting customer relationship information** in your firm's data warehouse. Please allocate 100 points across the following two situations to indicate **the relative frequency that you encounter these situations** . For example, if 80 percent of your **searches/requests** are for information in situations where you "react to a problem", please, place an 80 beside such situation. The total should be 100. **How often do your searches/requests involve information...**

(i) on **unknown problems/key issues** ? [ ] %

(ii) to help you **react** to a problem? [ ] %

Total 100%

50. Please allocate 100 points across the following two searching situations to indicate **the ratio of your time spent** in a typical week in such situations. For example, if 80 percent of your **time** searching for information it is spent "reacting", please, place an 80 beside such a situation. The total should be 100. **How much time do you spend...**

(i) on data warehouse **proactive** searches/requests without a particular problem to be addressed (e.g. monitoring a campaign)?. [ ] %

(ii) on data warehouse **reactive** searches/requests looking for information specific to a problem? [ ] %

Total 100%

51. Think of the 'rules of thumb' that you have applied in the last 6 months in order to simplify your search/request for information in your firm's data warehouse when you need to understand a customer relationship problem. Please, allocate 100 points across the following five information search/request strategies to indicate **their relative frequency** . For example, if in 40 percent of your **searches/requests** you use "The best problem template at hand", please, place a 40 beside such strategy. The total should be 100. **How often do your searches/requests involve the...**

(i) "The best 'problem template' at hand" ?: Looking for **common information** between the situation that you have to understand and one or several **'problem template(s)'** that you have at hand		%
ii) "The last occurrence" ?: Looking for information you used last time in this type of situation.		%
(iii) "The trial and error" ?: **Adjusting** your next question with the information that you got in a previous question		%
(iv) "The confirmatory" ?: Looking for information that **confirms** your belief about the problem		%
(v) Other information search/request strategies ?:		%
Total	100%	

52. Please allocate 100 points across the following five searching "rules of thumb" to indicate **the ratio of your time spent** in a typical week using such strategies. For example, if 80 percent of your **time** searching for information is spent using "the trial and error" strategy, please, place a 80 beside such strategy. The total should be 100. **How much time do you spend...**

(i) Using "The best 'problem template' at hand" strategy?		%
(ii) Using "The last occurrence" strategy?		%
(iii) Using "The trial and error" strategy?		%
(iv) Using "The confirmatory" strategy?		%
(v) Using other information search/request strategies?		%
Total	100%	

53. Think of the 'rules of thumb' that you have applied in the last 6 months in order to simplify your search/request for information in your firm's data warehouse when you need to understand a customer relationship problem. Please, allocate 100 points across the following five information search/request strategies to indicate **their relative frequency.** For example, if in a 60 percent of your **searches/requests** you use "The diagnostic", please, place a 60 beside such strategy. The total should be 100. **How often do your searches/requests involve the...**

(i) "The test" ?: Looking for information in order to identify an acceptable **sample/target** in order to be able to generalize to a larger scale		%
(ii) "The 'canned' queries" ?: Using **predefined** queries.		%
(iii) "The divide & conquer" ?: **Adding** successive selection **criteria** to your initial question		%
(iv) "The diagnostic" ?: Looking for information that **confirms** known symptoms.		%
(v) Other information search/request strategies ?:		%
Total	100%	

54. Again, please allocate 100 points across the following four searching "rules of thumb" to indicate **the ratio of your time spent** in a typical week using such strategies. For example, if 80 percent of your **time** searching for information it is spent using "the 'canned' queries" strategy, please, place a 80 beside such strategy. The total should be 100. **How much time do you spend...**

(i) Using "The test" search/request strategy?		%
(ii) Using "The 'canned' queries" search/request strategy?		%
(iii) Using "The divide and conquer" search/request strategy?		%
(iv) Using "The diagnostic" search/request strategy?		%
(v) Using other information search/request strategies?		%
Total	100%	

(iii) Using "The divide and conquer" search/request strategy? ____ %

(iv) Using "The diagnostic" search/request strategy? ____ %

(v) Using other information search/request strategies? ____ %

Total 100%

RESPONDENT'S PERSONAL DATA

The following information will only be used to provide an overall picture of the total sample. All data will be retained in the strictest confidence.

	Less than 6 months	Between 6 months and 1 year	Between 1 and 3 years	Between 3 and 5 years	More than 5 years
55. **How long** have you been involved in customer relationship management tasks?	○	○	○	○	○
56. How long have you been using your **company's data warehouse customer relationship data** to help you understand issues?	○	○	○	○	○
57. **How long** has your data warehouse been supporting CRM functions (any of them)?	○	○	○	○	○

	No time at all										All of my time
	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
58. In a **typical week** how much time do you spend **understanding CRM issues** (e.g. challenges, problems)	○	○	○	○	○	○	○	○	○	○	○
59. In a **typical week** how much time do you spend **searching/requesting customer information** from your firm's data warehouse?	○	○	○	○	○	○	○	○	○	○	○

59. In a **typical week** how much time do you spend **searching/requesting customer information** from your firm's data warehouse? ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○

Sales Support = sales analysis, sales history, contact management, lead tracking, etc.

Marketing Support = customer analysis, channel analysis, channel management, campaign management, etc.

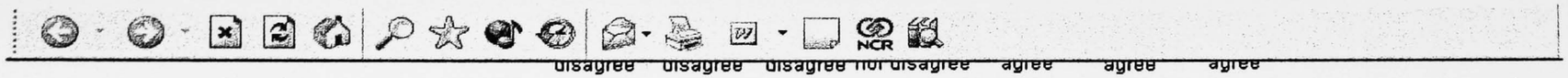
Customer Service Support = Call management, problem tracking, customer history, etc.

	Sales support	Marketing support	Customer service support	Other CRM functions
60. Which CRM functions are supported by your data warehouse? Check all that apply.	☐	☐	☐	☐
61. What is the **primary** CRM function of **your personal** job?.	○	○	○	○

To what extent do you agree or disagree with the following statements.

	Strongly disagree	Moderately disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Moderately agree	Strongly agree
62. My firm's data warehouse is an **important source** of customer relationship data for my personal job.	○	○	○	○	○	○	○
63. My firm's CRM activities target **individuals** as customers (i.e., not just other businesses or channel organizations)	○	○	○	○	○	○	○

Name (optional)



	disagree	disagree	disagree	nor disagree	agree	agree	agree
62. My firm's data warehouse is an **important source** of customer relationship data for my personal job.	○	○	○	○	○	○	○
63. My firm's CRM activities target **individuals** as customers (i.e., not just other businesses or channel organizations)	○	○	○	○	○	○	○

Name (optional)

Organization / Company (optional)

Would you like to receive a summary of the research findings? ○ Yes ○ No

If "YES"

Telephone contact

e-mail address

THANK YOU, FOR TAKING THE TIME TO RESPOND TO THIS SURVEY.
Please scroll down to the bottom of the page and click on the submit button.

Submit Reset

Created with HTML+Forms

APPENDIX F: DETAILS ON THE OPERATIONALIZATION OF VARIABLES

This research required five successive operationalizations. Each of the first four is linked to a questionnaire version and the fifth and last one is linked to the subsequent analysis made after factor analysis. Also, it has to be noted that only operationalizations #1 and #2 were never executed as questionnaire.v1 and questionnaire.v2 were not used in data collection. The following guidelines should be considered reading the details that it follows:

- Variables and items are qualified with a suffix .vn where n indicates the operationalization identifier. For example CUST.v1 indicates the operationalization #1 of the variable CUST which is linked to Questionnaire.v1.
- Labels of items should not be considered equivalent cross-operationalization. It means that, for example item_14.v1 and item_14_v2 might be two different items due to re-serialization, text modification, response format change, etc.
- Code reversed items are easily identifiable by the transformation. For example, 8-item_27.v1
- Analysis of the operationalization of a construct cross-operationalizations can be done by comparing the number of items. If needed, it might be convenient to see the items for a given construct through the four questionnaires.
- Operationalization #5 is made referencing as much as possible operationalization #4. In this way, differences are visible.
- Reliability measures are available for operationalizations #4 and #5 (and the original scales).

OPERATIONALIZATION OF VARIABLES IN QUESTIONNAIRE.v1

CUST.v1 = MEAN(item_10.v1,item_11.v1, item_12.v1, item_13.v1, Item_14.v1, item_15.v1)
COMP.v1 = MEAN(item_16.v1,item_17.v1)
TASKUT.v1 = MEAN(CUST.v1, COMP.v1)
ENAC.v1 = MEAN(item_18.v1,item_19.v1,item_20.v1,item_21.v1)
INT.v1 = MEAN(item_22.v1,item_23.v1,item_24.v1,item_25.v1,item_26.v1)
RICH.v1 = MEAN(item_30.v1, item_32.v1,item_33.v1,item_34.v1,item_37.v1)
EQUI.v1 = MEAN(8-item_27.v1,8-item_29.v1,8-item_36.v1)
TRUST.v1 = MEAN(8-item_28.v1,8-item_31.v1, item_35.v1,item_38.v1, item_39.v1)
ENACUT.v1 = MEAN(INT.v1, RICH.v1, EQUI.v1, ENAC.v1, TRUST.v1)
QUALITY.v1 = MEAN(TASKUT.v1, ENACUT.v1)

DKNOWE.v1 = MEAN(item_40.v1,item_41.v1,item_42.v1,item_43.v1)
PKNOWE.v1 = MEAN(item_44.v1, Item_45.v1,item_46.v1,item_47.v1)
ICPLXE.v1 = MEAN(item_48.v1,item_49.v1)
KNOWENH.v1 = MEAN(DKNOWE.v1, PKNOWE.v1, ICPLXE.v1)
LSEFE.v1 = MEAN(item_50.v1,item_51.v1,item_52.v1,item_53.v1)
MSELFE .v1 = MEAN(item_54.v1,item_55.v1,item_56.v1,item_57.v1)
HSELFE.v1 = MEAN(item_58.v1,item_59.v1,item_60.v1,item_61.v1)
SELFE.v1 = MEAN(LSEFE.v1, MSELFE .v1, HSELFE.v1)
ENACENH.v1= MEAN(KNOWENH.v1, SELFE.v1)

SCAN.v1 = MEAN(Tranformation$_{100}$(MEAN(item_62.v1,item_63v1,item_64.v1)) + item_67.v1)
FOCUS.v1 = MEAN(Tranformation$_{100}$(MEAN(item_65.v1,item_66.v1)) + item_68.v1)
REP.v1 = MEAN(item_69.v1, item_70.v1, item_77.v1)
AVBLE.v1 = MEAN(item_71.v1, item_72.v1,item_78.v1)
AA.v1 = MEAN(item_73.v1, item_74.v1,item_79.v1)
POSI.v1 = MEAN(item_75.v1,item_76.v1,item_80.v1)

OPERATIONALIZATION OF VARIABLES IN QUESTIONNAIRE.v2

CUST.v2 = MEAN(item_10.v2,item_11.v2, 8-item_12.v2, item_13.v2, Item_14.v2,8-item_15.v2)
COMP.v2 = MEAN(8-item_16.v2,item_17.v2)
TASKUT.v2 = MEAN(CUST.v2, COMP.v2)
ENAC.v2 = MEAN(item_18.v2,item_19.v2,item_20.v2,8-item_21.v2)
INT.v2 = MEAN(item_22.v2,item_23.v2,8-item_24.v2,8-item_25.v2,item_26.v2)
RICH.v2 = MEAN(item_30.v2, item_32.v2,8-item_33.v2,item_34.v2,item_37.v2)
EQUI.v2 = MEAN(8-item_27.v2,8-item_29.v2,8-item_36.v2)
TRUST.v2 = MEAN(8-item_28.v2,8-item_31.v2, item_35.v2,item_38.v2, 8-item_39.v2)
ENACUT.v2 = MEAN(INT.v2, RICH.v2, EQUI.v2, ENAC.v2, TRUST.v2)
QUALITY.v2 = MEAN(TASKUT.v2, ENACUT.v2)

DKNOWE.v2 = MEAN(item_40.v2,item_41.v2,8-item_42.v2,item_43.v2)
PKNOWE.v2 = MEAN(item_44.v2, item_45.v2,item_46.v2,8-item_47.v2)
ICPLXE.v2 = MEAN(item_48.v2,8-item_49.v2)
KNOWENH.v2 = MEAN(DKNOWE.v2, PKNOWE.v2, ICPLXE.v2)
LSEFE.v2 = MEAN(item_50.v2,item_51.v2,item_52.v2,item_53.v2)
MSELFE .v2 = MEAN(item_54.v2,item_55.v2,item_56.v2,item_57.v2)
HSELFE.v2 = MEAN(item_58.v2,item_59.v2,item_60.v2,item_61.v2)
SELFE.v2 = MEAN(LSEFE.v2, MSELFE .v2, HSELFE.v2)
ENACENH.v2= MEAN(KNOWENH.v2, SELFE.v2)

SCAN.v2 = MEAN(item_62i.v2,item_63iv1, item_64i.v2)
FOCUS.v2 = MEAN(item_62ii.v2,item_63ii.v2, item_64ii.v2)
REP.v2 = MEAN(item_65i.v2, item_66i.v2, item_67i.v2,item_68i.v2)
AVBLE.v2 = MEAN(item_65ii.v2, item_66ii.v2, item_67ii.v2,item_68ii.v2)
AA.v2 = MEAN(item_65iii.v2, item_66iii.v2, item_67iii.v2,item_68iii.v2)
POSI.v2 = MEAN(item_65iv.v2, item_66iv.v2, item_67iv.v2,item_68iv.v2)

OPERATIONALIZATION OF VARIABLES IN QUESTIONNAIRE.v3

CUST.v3 = MEAN(item_1.v3,item_2.v3, 8-item_3.v3, item_4.v3, Item_5.v3, 8-item_6.v3)
COMP.v3 = MEAN(8-item_7.v3,item_8.v3)
TASKUT.v3 = MEAN(CUST.v3, COMP.v3)
ENAC.v3 = MEAN(item_9.v3,item_10.v3,item_11v3,8-item_12.v3)
INT.v3 = MEAN(item_13.v3,item_14.v3,8-item_15.v3,8-item_16.v3,item_17.v3)
RICH.v3 = MEAN(item_21.v3, item_23.v3, 8-item_24.v3, item_25.v3,item_28.v3)
EQUI.v3 = MEAN(8-item_18.v3,8-item_20.v3,8-item_27.v3)
TRUST.v3 = MEAN(8-item_19.v3,8-item_22.v3, item_26.v3, item_29.v3,8-item30.v3)
ENACUT.v3 = MEAN(INT.v3, RICH.v3, EQUI.v3, ENAC.v3, TRUST.v3)
QUALITY.v3 = MEAN(TASKUT.v3, ENACUT.v3)

DKNOWE.v3 = MEAN(item_31.v3,item_32.v3,8-item_33.v3,item_34.v3)
PKNOWE.v3 = MEAN(item_35.v3, item_36.v3,item_37.v3,8-item_38.v3)
ICPLXE.v3 = MEAN(item_39.v3,8-item_40.v3)
KNOWENH.v3 = MEAN(DKNOWE.v3, PKNOWE.v3, ICPLXE.v3)
LSEFE.v3 = MEAN(item_41.v3,item_42.v3,item_43.v3,item_44.v3)
MSELFE .v3 = MEAN(item_45.v3,item_46.v3,item_47.v3,item_48.v3)
HSELFE.v3 = MEAN(item_49.v3,item_50.v3,item_51.v3,item_52.v3)
SELFE.v3 = MEAN(LSEFE.v3, MSELFE .v3, HSELFE.v3)
ENACENH.v3= MEAN(KNOWENH.v3, SELFE.v3)

SCAN.v3 = MEAN(item_53i.v3,item_54i.v3,item_55i.v3)
FOCUS.v3 = MEAN(item_53ii.v3,item_54ii.v3,item_55ii.v3)
REP.v3 = MEAN(item_56i.v3, item_57i.v3, item_58i.v3,item_59i.v3)
AVBLE.v3 = MEAN(item_56ii.v3, item_57ii.v3,item_58ii.v3,item_59ii.v3)
AA.v3 = MEAN(item_56iii.v3, item_57iii.v3,item_58iii.v3,item_59iii.v3)
POSI.v3 = MEAN(item_56iv.v3,item_57iv.v3,item_58iv.v3, item_59iv.v3)

OPERATIONALIZATION OF VARIABLES IN QUESTIONNAIRE.v4

CUST.v4 = MEAN(item_1.v4,item_2.v4, 8-item_3.v4, item_4.v4, Item_5.v4, 8-item_6.v4)
COMP.v4 = MEAN(item_7.v4,item_8.v4,item_9.v4)
TASKUT.v4 = MEAN(CUST.v4, COMP.v4)
ENAC.v4 = MEAN(item_10.v4,item_11.v4,item_12.v4,8-item_13.v4)
INT.v4 = MEAN(item_14.v4,item_15.v4,item_16.v4)
RICH.v4 = MEAN(item_20.v4, item_22.v4,8-item_23.v4,item_24.v4,item_27.v4)
EQUI.v4 = MEAN(8-item_17.v4,8-item_19.v4,8-item_26.v4)
TRUST.v4 = MEAN(item_18.v4,8-item_21.v4, item_25.v4)
ENACUT.v4 = MEAN(INT.v4, RICH.v4, EQUI.v4, ENAC.v4, TRUST.v4)
QUALITY.v4 = MEAN(TASKUT.v4, ENACUT.v4)

DKNOWE.v4 = MEAN(item_28.v4,item_29.v4,item_30.v4)
PKNOWE.v4 = MEAN(item_31.v4, item_32.v4,item_33.v4,item_34.v4)
ICPLXE.v4 = MEAN(item_35.v4,item_36.v4)
KNOWENH.v4 = MEAN(DKNOWE.v4, PKNOWE.v4, ICPLXE.v4)
LSEFE.v4 = MEAN(item_37.v4,item_38.v4,item_39.v4,item_40.v4)
MSELFE .v4 = MEAN(item_41.v4,item_42.v4,item_43.v4,item_44.v4)
HSELFE.v4 = MEAN(item_45.v4,item_46.v4,item_47.v4,item_48.v4)
SELFE.v4 = MEAN(LSEFE.v4, MSELFE .v4, HSELFE.v4)
ENACENH.v4= MEAN(KNOWENH.v4, SELFE.v4)

SCAN.v4 = MEAN(item_49i.v4,item_50i.v4)
FOCUS.v4 v4 = MEAN(item_49ii.v4,item_50ii.v4)
REP.v4 = MEAN(item_51i.v4, item_52i.v4, item_53i.v4,item_54i.v4)
AVBLE.v4 = MEAN(item_51ii.v4, item_52ii.v4,item_53ii.v4,item_54ii.v4)
AA.v4 = MEAN(item_51iii.v4, item_52iii.v4,item_53iii.v4,item_54iii.v4)
POSI.v4 = MEAN(item_51iv.v4,item_52iv.v4,item_53iv.v4, item_54iv.v4)

OPERATIONALIZATION OF VARIABLES AFTER FACTOR ANALYSIS

CUST.v5 = MEAN(item_1.v4,item_2.v4)
COMP.v5 = COMP.v4
TASKUT.v5 = MEAN(CUST.v5, COMP.v4)
INT.v5 = MEAN(8-item_3.v4,item_15.v4,item_24.v4)
RICH.v5 = MEAN(item_14.v4,item_18.v4, item_20.v4, item_22.v4, item_25.v4)
EQUI.v5 = MEAN(8-item_17.v4,8-item_26.v4)
ENACUT.v5 = MEAN(INT.v5, RICH.v5, EQUI.v5)
QUALITY.v5 = MEAN(TASKUT.v5, ENACUT.v5)

KNOWENH.v5 = MEAN(item_28.v4,item_29.v4,item_30.v4, item_32.v4, item_34.v4, item_35.v4, item_36.v4)
LSEFE.v5 = LSEFE.v4
MSELFE .v5 = MSELFE .v4
HSELFE.v5 = HSELFE.v4
SELFE.v5 = SELFE.v4
ENACENH.v5= MEAN(KNOWENH.v5, SELFE.v4)

SCAN.v5 = SCAN.v4
FOCUS.v5 = FOCUS.v4
REP.v5 = REP.v4
AVBLE.v5 = AVBLE.v4
AA.v5 = AA.v4
POSI.v5 = POSI.v4

APPENDIX G: CALLS FOR PARTICIPATION SUPPORT

CALL FOR PARTICIPATION TO SEGMENT 21&22. MODE: E-MAIL

Dear (%FIRST NAME%),

As a research associate at Henley Management College, in the UK, I am conducting an international academic not-for-profit research project on Customer Relationship Management supported by a data warehouse. By data warehouse it is meant a central repository of customer data from diverse sources. The objectives of the research are to reach an accurate understanding of

- CRM practitioners' perceptions of the quality of customer relationship data in their firms' data warehouse
- The impact of a data warehouse' customer relationship data on CRM practitioners' ability understanding CRM problems, e.g., identification of the components of a customer relationship problem statement
- The impact of a data warehouse customer relationship data on CRM practitioners' information search behaviour, e.g., search strategies.

Key Points

In my experience, working in the area of marketing information systems for 25 years as a practitioner, key points are that

- CRM practitioners' perceptions of the quality of customer relationship data are a key factor in their confidence using data and making decisions. Such perceptions might result in a gap with respect some metrics that the IT executives might have (e.g. using data profiling tools). Remember: "Perception is reality"
- Today's competitive pressure requires permanent justification of your business assets. Wouldn't be great to demonstrate that thanks to your data warehouse CRM practitioners have improved (i) their ability understanding CRM problems and (ii) their information search behaviour?

Invitation to Participate

(%FIRST NAME%), I need qualified informants (i.e. **practitioners working in campaigns, sales, or customer support with customer data stored in a data warehouse)** willing to participate in this research. I kindly ask for your participation filling out the survey and help forwarding the URL below to other qualified informants. Please, notice that it is NOT required to have in production a CRM application (i.e. end-users might access customer data via traditional queries and/or business intelligence tools).

Value Proposition for You

By filling out the short survey (20 minutes) at the URL below,

- You will be able to reuse this scientific questionnaire for assessing periodically your CRM practitioners' perceptions of the quality of customer relationship data. Doing it, you will be able to track evolution and plan/implement corrective actions.
- You will receive a summary of the results

- (Optional) you can benchmark your organization to the overall findings.

Action to be Taken

- Please, fill out the short survey at the URL
www.henleymc.ac.uk/quest/59365.htm
- Forward the URL www.henleymc.ac.uk/quest/59365.htm to your **CRM practitioners working with customer data in campaigns, sales, or customer support.**
- In case you want to benchmark your organization, you should (i) ask your CRM practitioners for fulfilling the organization name with a coded name (e.g. "Bank of Joe"), an agreed acronym or the explicit true name and (ii) forward to my attention the e-mail address of the person designated for receiving the benchmark. Please, notice that identification (e.g. personal name, organization) is optional.

Follow Up

- If we get a significant response rate from your organization (i.e. +10 respondents), we could benchmark the results to the overall findings and send a report to the attention of someone in your organization.
- We will send a summary of the results if this is requested in the questionnaire.
- The questionnaire will be available for you in case you request it

(%FIRST NAME%), I hope you will find interesting my kind invitation to participate in this research. Otherwise, please, accept my apologies for disturbing you. If you have questions in regard to this research please contact me. If you require further confirmation of my status as Research Associate of Henley Management College, please, contact the Director of Studies, Doctoral Programme, Dr. David Price (David.Price@henleymc.ac.uk).

Sincerely,

Raul M. Abril
HMC, Research Associate
Tel. USA: +1 760 233 08 29
ra.dba15@henleymc.ac.uk

CALL FOR PARTICIPATION TO SEGMENT 21&22. MODE: LETTERHEAD

Dear (%FIRST NAME%),

I am a research associate at Henley Management College conducting an international academic not-for-profit research project. The research focuses on the quality of data stored in data warehouses, and how this supports customer relationship management (data warehouse is defined as a central repository of customer data from diverse sources). It is NOT required to have in production a CRM application (i.e,. end-users might access customer data via traditional queries and/or business intelligence tools). The objectives of the research are to allow data warehouse practitioners and marketing managers to reach an accurate understanding of the following key points:

- *End-users' perceptions of the quality of customer relationship data in their firms' data warehouse.* Such perceptions might result in a gap with respect to some metrics that data warehouse practitioners might obtain from data profiling tools.
- *The impact of data warehouse' customer relationship data on end-users' (i) ability to understand CRM problems and (ii) search behaviour.* Wouldn't it be great to demonstrate that thanks to the data warehouse end-users have improved (i) their ability to understand problems and (ii) their information search behaviour?

Invitation to Participate
(%FIRST NAME%), I need qualified informants (i.e., practitioners working in campaigns, sales, or customer support with customer data stored in a data warehouse) willing to participate in this research. I kindly ask for (i) your participation in filling out the survey, and (ii) help forwarding the URL for the survey to other qualified informants.

Value Proposition for You
By filling out the short survey (20 minutes) at the URL below,
- You will receive a summary of the results.
- You will be able to reuse this scientific questionnaire for assessing periodically your CRM end-users' perceptions of the quality of your customer relationship data. Doing so, you will be able to track evolution and plan/implement corrective actions.
- You can benchmark your organization to the overall findings. If we get a significant response rate from your organization (i.e. >10 respondents), we could benchmark the results to the overall findings and send a report to the attention of someone in your organization.

Action to be Taken
- Forward the URL www.henleymc.ac.uk/quest/59365.htm to your data warehouse end-users working with customer data in campaigns, sales, or customer support.
- In case you want to benchmark your organization, you should (i) ask your CRM end-users to complete the organization name with a coded name (e.g. "Bank of Joe"), or the explicit true name and (ii) forward to my attention the e-mail address of the person designated for receiving the benchmark. Please, notice that identification (e.g. personal name, organization) is optional.

(%FIRST NAME%), if you have questions in regard to this research please contact me. If you require further confirmation of my status as Research Associate of Henley Management College, please contact the Director of Studies, Doctoral Programme, Dr. David Price (David.Price@henleymc.ac.uk).

Sincerely,

Raul M. Abril
HMC, Research Associate
Tel. USA: +1 760 233 08 29
ra.dba15@henleymc.ac.uk

FOLLOW UP CALL FOR SUPPORT TO SEGMENT 11&12&C.
MODE: E-MAIL

Dear (%FIRST NAME%),

Last (%DATE%), I sent a letter to your attention (the text is attached). I got your address in a conference where I was one of the speakers. I am not sure if you got it. So, I wanted to send this via e-mail. Hopefully, this will reach you.

Best regards.

Raul M. Abril
HMC, Research Associate
ra.dba15@henleymc.ac.uk

–

Dear (%FIRST NAME%),

I am a research associate at Henley Management College conducting an international academic not-for-profit research project. The research focuses on the quality of data stored in data warehouses, and how this supports customer relationship management (data warehouse is defined as a central repository of customer data from diverse sources). It is NOT required to have in production a CRM application (i.e,. end-users might access customer data via traditional queries and/or business intelligence tools). The objectives of the research are to allow data warehouse practitioners and marketing managers to reach an accurate understanding of the following key points:

- *End-users' perceptions of the quality of customer relationship data in their firms' data warehouse.* Such perceptions might result in a gap with respect to some metrics that data warehouse practitioners might obtain from data profiling tools.
- *The impact of data warehouse' customer relationship data on end-users' (i) ability to understand CRM problems and (ii) search behaviour.* Wouldn't it be great to demonstrate that thanks to the data warehouse end-users have improved (i) their ability to understand problems and (ii) their information search behaviour?

Invitation to Participate
(%FIRST NAME%), I need qualified informants (i.e., practitioners working in campaigns, sales, or customer support with customer data stored in a data warehouse) willing to participate in this research. I kindly ask for (i) your participation in filling out the survey, and (ii) help forwarding the URL for the survey to other qualified informants.

Value Proposition for You
By filling out the short survey (20 minutes) at the URL below,
- You will receive a summary of the results.

- You will be able to reuse this scientific questionnaire for assessing periodically your CRM end-users' perceptions of the quality of your customer relationship data. Doing so, you will be able to track evolution and plan/implement corrective actions.
- You can benchmark your organization to the overall findings. If we get a significant response rate from your organization (i.e. >10 respondents), we could benchmark the results to the overall findings and send a report to the attention of someone in your organization.

Action to be Taken

- Forward the URL www.henleymc.ac.uk/quest/59365.htm to your data warehouse end-users working with customer data in campaigns, sales, or customer support.
- In case you want to benchmark your organization, you should (i) ask your CRM end-users to complete the organization name with a coded name (e.g. "Bank of Joe"), or the explicit true name and (ii) forward to my attention the e-mail address of the person designated for receiving the benchmark. Please, notice that identification (e.g. personal name, organization) is optional.

(%FIRST NAME%), if you have questions in regard to this research please contact me. If you require further confirmation of my status as Research Associate of Henley Management College, please contact the Director of Studies, Doctoral Programme, Dr. David Price (David.Price@henleymc.ac.uk).

Sincerely,

Raul M. Abril
HMC, Research Associate
Tel. USA: +1 760 233 08 29
ra.dba15@henleymc.ac.uk

CALL FOR SUPPORT TO SEGMENT 31. MODE: E-MAIL

Dear (%FIRST NAME%),

I am approaching you as the contact person in the (%SOURCE%). I am a research associate at Henley Management College conducting an international academic not-for-profit research project. The research focuses on the quality of data stored in data warehouses, and how this supports customer relationship management (data warehouse is defined as a central repository of customer data from diverse sources). It is NOT required to have in production a CRM application (i.e,. end-users might access customer data via traditional queries and/or business intelligence tools). The objectives of the research are to allow data warehouse practitioners and marketing managers to reach an accurate understanding of the following key points:

1. *End-users' perceptions of the quality of customer relationship data in their firms' data warehouse.* Such perceptions might result in a gap with respect to some metrics that data warehouse practitioners might obtain from data profiling tools.
2. *The impact of data warehouse' customer relationship data on end-users' (i) ability to understand CRM problems and (ii) search behaviour.* Wouldn't it be great to demonstrate that thanks to the data warehouse end-users have improved (i) their ability to understand problems and (ii) their information search behaviour?

Invitation to Participate

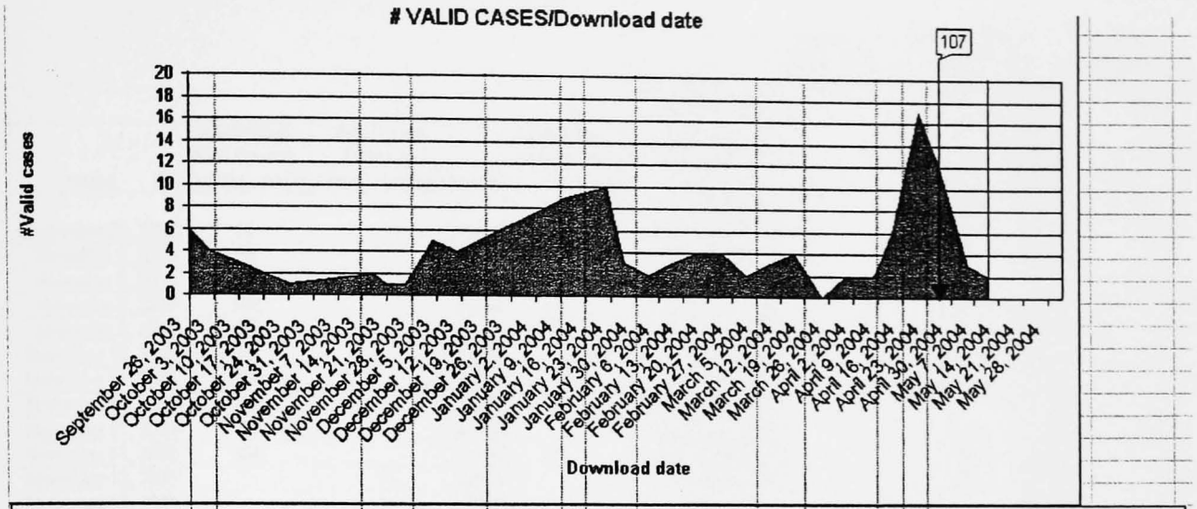
(%FIRST NAME%), I need qualified informants (i.e., end-users working in campaigns, sales, or customer support with customer data stored in a data warehouse) willing to participate in this research. **I kindly ask for your help forwarding the URL** www.henleymc.ac.uk/quest/59365.htm **for the survey to the members of the (%SOURCE%) chapter asking for their collaboration finding qualified informants in their organizations.**

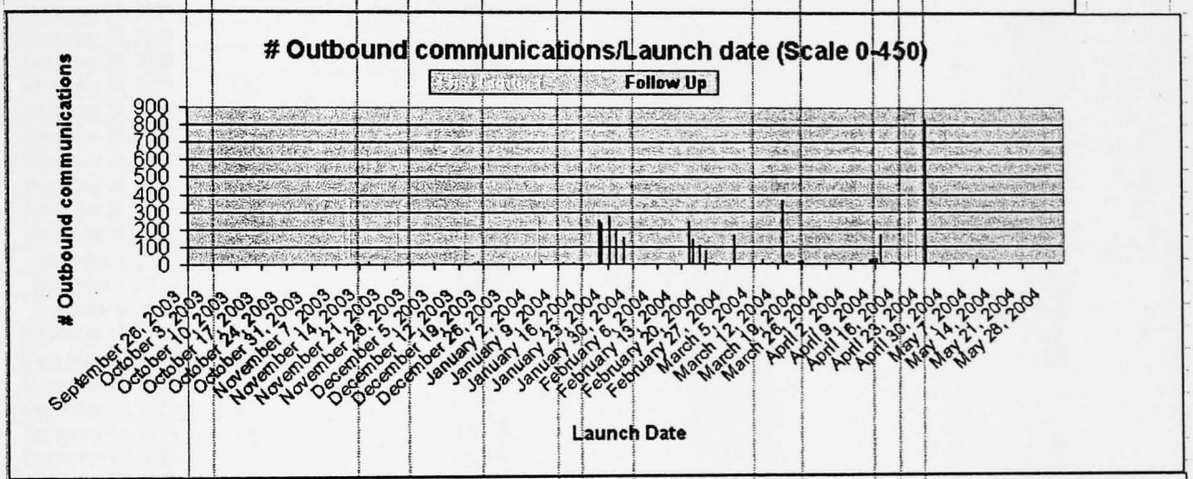
(%FIRST NAME%), if you have questions in regard to this research please contact me. If you require further confirmation of my status as Research Associate of Henley Management College, please contact the Director of Studies, Doctoral Programme, Dr. David Price (David.Price@henleymc.ac.uk).

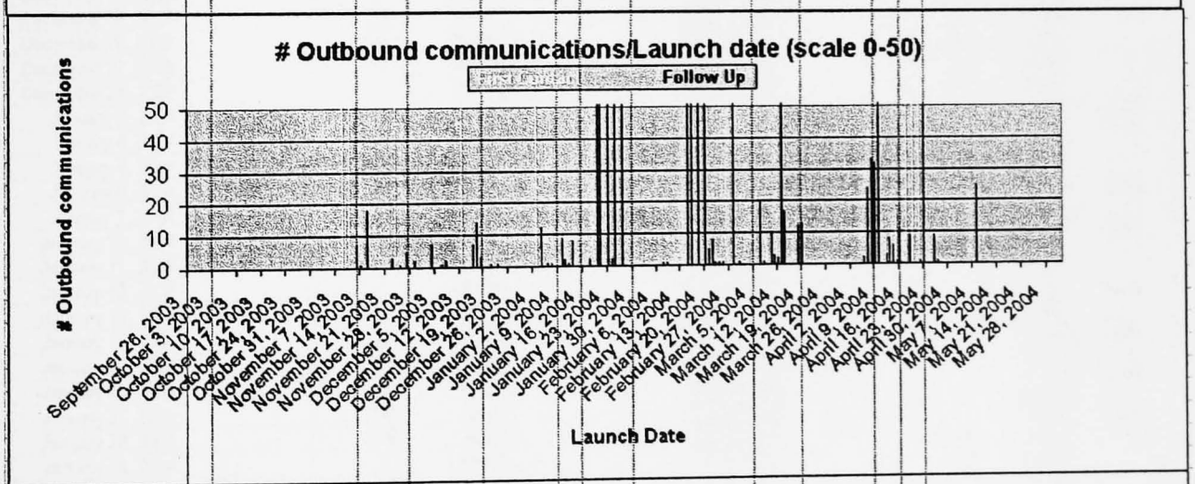
Sincerely,

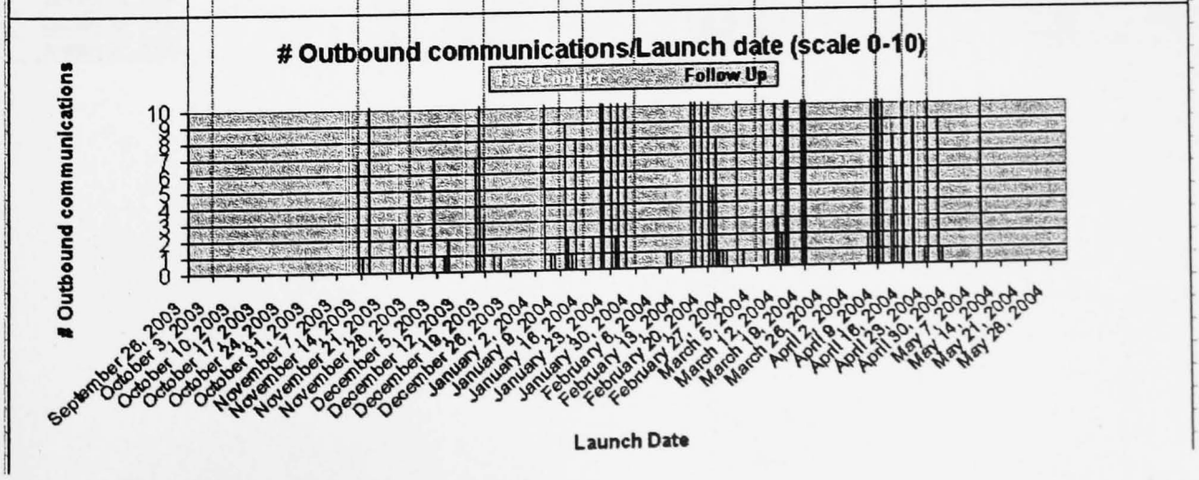
Raul M. Abril
HMC, Research Associate
Tel. USA: +1 760 233 08 29
ra.dba15@henleymc.ac.uk

APPENDIX H: OUTBOUND COMMUNICATIONS vs. VALID CASES CHARTS





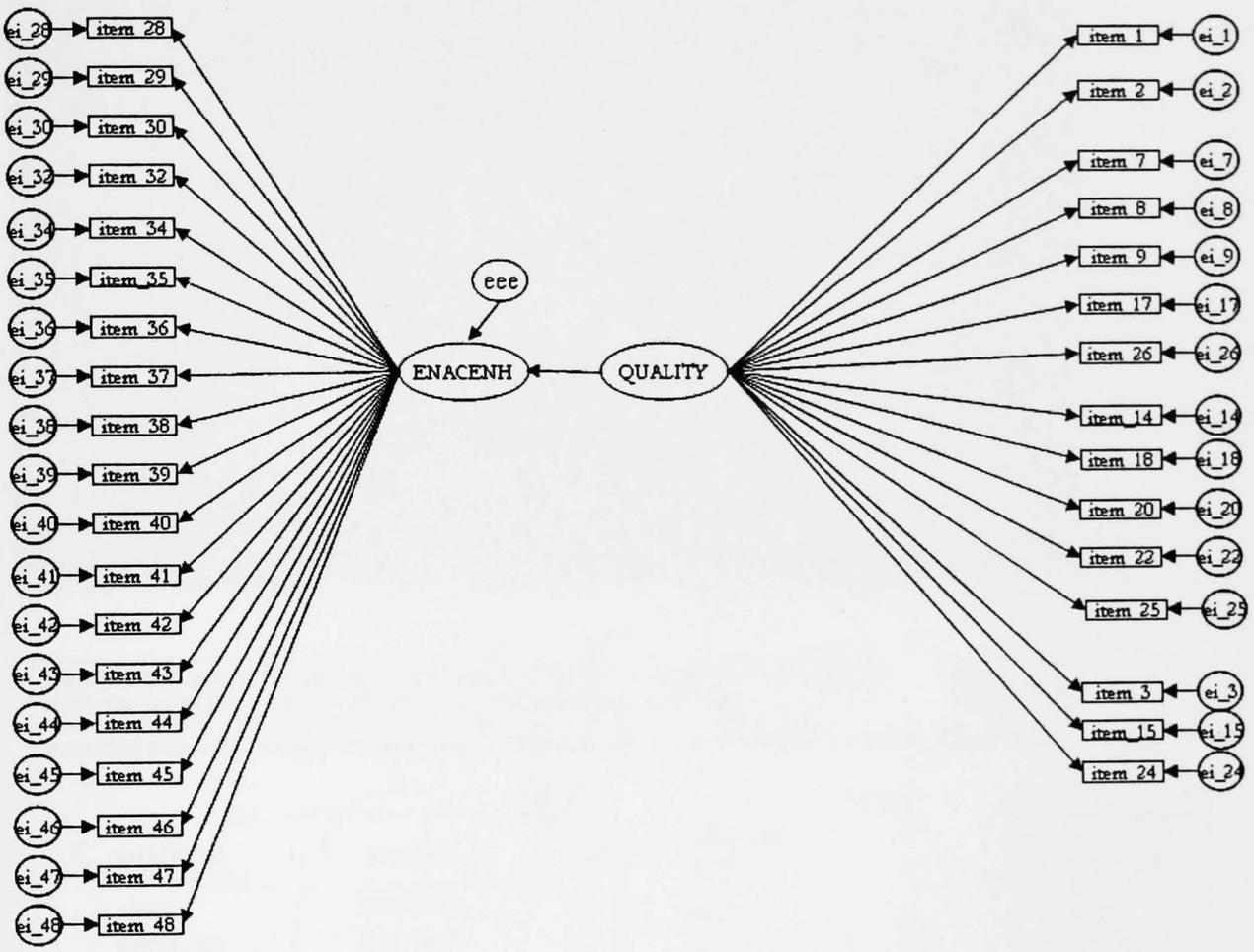




APPENDIX I: MULTIMODE-OUTBOUND COMMUNICATIONS

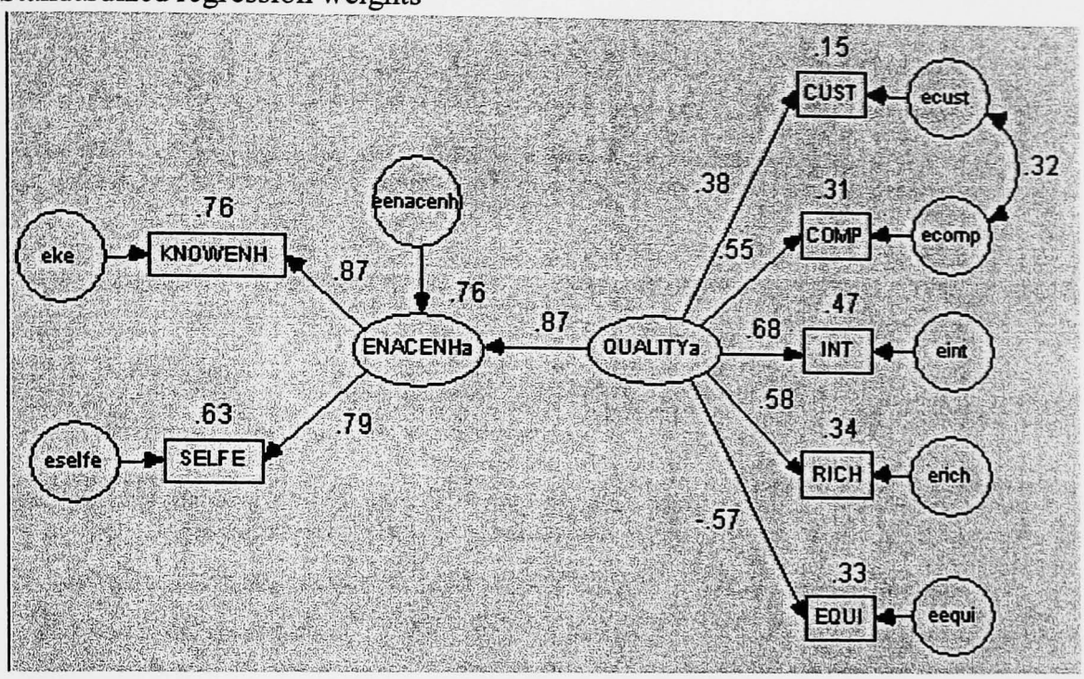
Period	# Outbound First Contact Coomunicatio	# Outbound Follow up Coomunicatior	Me	Leeds in existing list? (L)
September 26, 2003	15		Paper	OWN
November 9, 2003	99		e-mail	L
November 9, 2003	14		e-mail	L
November 9, 2003	198		e-mail	L
November 9, 2003	85		e-mail	L
November 10, 2003	22		e-mail	L
November 11, 2003	372		e-mail	L
November 11, 2003	119		e-mail	L
November 11, 2003	36		e-mail	L
November 11, 2003	364		e-mail	L
November 13, 2003		7	e-mail	
November 14, 2003		1	e-mail	
November 16, 2003		18	e-mail	
November 23, 2003		3	e-mail	
November 24, 2003	130		e-mail	L
November 24, 2003	16		e-mail	L
November 24, 2003	15		e-mail	L
November 24, 2003	314		e-mail	L
November 25, 2003		1	e-mail	
November 27, 2003		5	e-mail	
November 29, 2003		2	e-mail	
December 4, 2003		7	e-mail	
December 7, 2003		1	e-mail	
December 8, 2003		2	e-mail	
December 13, 2003	186		e-mail	OWN
December 13, 2003	307		e-mail	L
December 13, 2003	24		e-mail	L
December 13, 2003	21		e-mail	L
December 13, 2003	21		e-mail	L
December 16, 2003		7	e-mail	
December 17, 2003		14	e-mail	
December 18, 2003		3	e-mail	
December 21, 2003		1	e-mail	
December 23, 2003		1	e-mail	
January 3, 2004	267		e-mail	OWN
January 4, 2004	5		e-mail	OWN
January 4, 2004		12	e-mail	
January 6, 2004		1	e-mail	
January 7, 2004		1	e-mail	
January 10, 2004		9	e-mail	
January 11, 2004		2	e-mail	
January 12, 2004		1	e-mail	
January 13, 2004		8	e-mail	
January 18, 2004		2	e-mail	
January 20, 2004		252	mail	
January 20, 2004		4	e-mail	
January 21, 2004		238	mail	
January 23, 2004		278	mail	
January 24, 2004		2	e-mail	
January 25, 2004		189	mail	
January 27, 2004		156	mail	
February 8, 2004		1	e-mail	
February 9, 2004		1	e-mail	
February 15, 2004		1	e-mail	
February 15, 2004		230	e-mail	
February 15, 2004		171	e-mail	
February 15, 2004		241	e-mail	
February 15, 2004		145	e-mail	
February 16, 2004		144	mail	
February 18, 2004		106	mail	
February 20, 2004		80	mail	
February 20, 2004		2	e-mail	
February 21, 2004		5	e-mail	
February 22, 2004		8	e-mail	
February 23, 2004		1	e-mail	
February 24, 2004		1	e-mail	
February 25, 2004		1	e-mail	
February 28, 2004		1	e-mail	
February 28, 2004		168	e-mail	
February 28, 2004		120	e-mail	
March 1, 2004		1	e-mail	
March 7, 2004		20	e-mail	
March 8, 2004		1	e-mail	
March 10, 2004		10	e-mail	
March 11, 2004		3	e-mail	
March 12, 2004		2	e-mail	
March 13, 2004		368	e-mail	
March 14, 2004		3	e-mail	
March 14, 2004		17	e-mail	
March 14, 2004		2	e-mail	
March 18, 2004		12	e-mail	
March 19, 2004		13	e-mail	
April 6, 2004		2	e-mail	
April 7, 2004		24	e-mail	
April 8, 2004		33	e-mail	
April 9, 2004		32	e-mail	
April 10, 2004		32	e-mail	
April 10, 2004		89	e-mail	
April 10, 2004		63	e-mail	
April 10, 2004		36	e-mail	
April 10, 2004		93	e-mail	
April 10, 2004		135	e-mail	
April 10, 2004		167	e-mail	
April 10, 2004		88	e-mail	
April 10, 2004		74	e-mail	
April 10, 2004		87	e-mail	
April 10, 2004		44	e-mail	
April 12, 2004	7		e-mail	OWN
April 12, 2004	9		e-mail	OWN
April 12, 2004		3	e-mail	
April 13, 2004		3	e-mail	
April 13, 2004	7		e-mail	OWN
April 13, 2004		8	e-mail	
April 13, 2004	3		e-mail	OWN
April 14, 2004		6	e-mail	
April 15, 2004		1	e-mail	
April 15, 2004	17		e-mail	OWN
April 16, 2004		11	e-mail	
April 18, 2004	414		e-mail	OWN
April 19, 2004		9	e-mail	
April 20, 2004	260		e-mail	OWN
April 22, 2004		1	e-mail	
April 22, 2004	170		e-mail	OWN
April 23, 2004		829	e-mail	
April 26, 2004		9	e-mail	
April 27, 2004		1	e-mail	
May 8, 2004		25	e-mail	

APPENDIX J: DETAILED MODEL TO BE TESTED USING STRUCTURAL EQUATION MODELING



APPENDIX K: REDUCED MODEL TESTED USING STRUCTURAL EQUATION MODELING

Standardized regression weights



The standardized regression weights represent the amount of change in the dependent variable that is attributable to a single standard deviation unit's worth of change in the predictor variable.

$O1 = b_1 L1 + eO1$
$L1 = b_1 I3 + eL1$

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
ENACENHa	<---	QUALITYa	.871
SELFE	<---	ENACENHa	.795
KNOWENH	<---	ENACENHa	.872
CUST	<---	QUALITYa	.384
COMP	<---	QUALITYa	.554
INT	<---	QUALITYa	.684
RICH	<---	QUALITYa	.582
EQUI	<---	QUALITYa	-.572

Predictors

Beta standardized regression coefficients are <1. When **ENACENH** goes up by 1 standard deviation, **SELFE** goes up by 0.795 standard deviations. They rank between high and moderate [Hair Jr., J. F. 2003 #3264]. CUST, COMP, INT, RICH and EQUI fall below the threshold of .7 [Hair Jr., J. F. 1998 #1759] pag 612. The rank .41 to .70 is considered moderate.

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
ENACENHa	.759
EQUI	.327
RICH	.339
INT	.468
COMP	.307
CUST	.148
KNOWENH	.760
SELFE	.632

Under estimate are the R2 values for each indicator (observed and latent) variable indicating the proportion of its variance explainable by its predictor. It is estimated that **QUALITY** (the predictor of **ENACENH**) explains 75.9 percent of its variance. In other words, the error variance of **ENACENH** is approximately 24.1 percent of the variance of **ENACENH** itself. R2 for EQUI is .327 = -.572 * -.572. INT falls slightly bellow .5. CUST,COMP, RICH and EQUI have a marginal level of acceptance.

Indicators' reliability is measured by R2. The indicators's reliability should exceed .50, which roughly corresponds to a beta standard regression coefficient of .70 [Hair Jr., J. F. 1998 #1759] pag 612.

Once the overall model fit has been evaluated, the measurement of each construct should be assessed For unidimensionality and reliability [Hair Jr., J. F. 1998 #1759]

Standardized Direct Effects (Group number 1 - Default model)

	QUALITYa	ENACENHa
ENACENHa	.871	.000
EQUI	-.572	.000
RICH	.582	.000
INT	.684	.000
COMP	.554	.000
CUST	.384	.000
KNOWENH	.000	.872
SELFE	.000	.795

$O1=b_1L1+eO1$
$L1=b_1I3+eL1$

Reliability of latent constructs in Confirmatory Factor Analysis can be assessed with two measures: The composite construct reliability and the Variance extracted [Hair Jr., J. F. 1998 #1759].
The composite construct reliability measures the internal consistency of the construct predictors, depicting the degree to which they "indicate" the common latent (unobserved) construct [Hair Jr., J. F. 1998 #1759]. A commonly used threshold value for acceptable reliability is .70 although values below .70 have been deemed acceptable if the research is exploratory in nature [Hair Jr., J. F. 1998 #1759].

The variance extracted R^2 by the latent construct is the amount of variance in its indicators that the latent construct is accounted for. A commonly used threshold value for acceptable reliability is .50 [Hair Jr., J. F. 1998 #1759].

QUALITY has a marginal composite construct reliability. However, ENACENH exceeds the threshold of .7 [Hair Jr., J. F. 1998 #1759]

	Composite Construct Reliability (Phase II)	Variance extracted (Phase II)
ENACENHa	0.82	0.70
QUALITYa	0.44	0.32

The variances extracted for ENACENH exceeds the threshold of .5 However in QUALITY this is not the case [Hair Jr., J. F. 1998 #1759].

Standardized Residual Covariances (Group number 1 - Default model)

	EQUI	RICH	INT	COMP	CUST	KNOWENH	SELFE
EQUI	.000						
RICH	-.118	.000					
INT	.242	.188	.000				
COMP	-.179	-1.105	.467	.000			
CUST	.180	-.749	.994	.000	.000		
KNOWENH	.170	.521	-.377	.280	-.057	.000	
SELFE	-.436	-.194	.057	-.239	-.453	.000	.000

All the standardized residuals are < 2.58. Therefore, there are not statistically significant discrepancies with zero residuals (i.e. perfect model fit)

No modification indices were displayed, this means that none exceed the specified threshold.

Model goodness-of-fit analysis

Type of Fit	Measure	Interpretation	Suggested Cut-off value	MODEL
Absolute	p of χ^2	Nonsignificant differences. The model fits the sample data	>.2	0.40
	p of Bollen-Stine bootstrap	Nonsignificant differences. The model fits the bootstrap sample data	>.05	0.17
	GFI	Relative amount of variance in the observed variables that is jointly explained by the estimated population variance.	≥.90	0.97
	RMR standardized	The average discrepancy per degree of freedom between the sample observed and hypothesized correlation matrices. It can be interpreted as meaning that the model explains the correlations to within an average error of <RMR>.	<.05	0.04
	RMSEA	Similar to the RMR, The RMSEA is the average difference per degree of freedom. It differs from the RMR, however, in that the discrepancy is measured in terms of the population, not just the sample used for estimation.	<.08	0.02
	HOELTER .05	If >200 then it suggest that the sample size is appropriate for the calculation of estimates	>200	177
	HOELTER .01	If >200 then it suggest that the sample size is appropriate for the calculation of estimates	>200	220
Incremetal	AGFI	It is the GFI adjusted by the ratio of degrees of freedom for the proposed model to the degrees of freedom for the null model. The null model is hypothesized to be the simplest model that can be theoretically justified	≥.90	0.93
	NFI	Relative comparison of the proposed model to the null model.	≥.90	0.95
Parsimonius	χ^2/d.f.	If < 1.0 then it suggests that there are too many coefficients, If > 2.0 or 3.0 it suggests that the model is not yet representative of the observed data	Lower limit:1.0 Upper limit:3:0	1.05

Given the statistical properties of ML and its characteristics at smaller sample sizes , the reseache is encouraged to be conservative in specifying a significance level, chossing smaller levels (.025 or .01) instead of the traditional .05 level [Hair Jr., J. F. 1998 #1759]. In the case of the χ2 the conservative significance levels will be big values (.1 or .2)

[Byrne, B. M. 2001 #3261][Hair Jr., J. F. 1998

[Byrne, B. M. 2001 #3261][Hair Jr., J. F. 1998 #1759]

[Hair Jr., J. F. 1998 #1759]

[Byrne, B. M. 2001 #3261]

[Byrne, B. M. 2001 #3261]

[Byrne, B. M. 2001 #3261][Hair Jr., J. F. 1998

[Byrne, B. M. 2001 #3261][Hair Jr., J. F. 1998 #1759]

[Hair Jr., J. F. 1998 #1759]
<1.0 "overfitted" model
> 2.0 or 3.0 model not yet representative of the observed data

APPENDIX L: ANOVA ANALYSES TESTING GROUPS DIFFERENCES ON VARIABLES INVOLVED IN ASSOCIATIONS

FACTOR: CRM job function (JOBFUNC)

Descriptive Statistics

Dependent Variable: The Quality of DW Customer Relationship Data for Problem Enactment

61 What is the primary	Mean	Std. Deviation	N
Sales Support	4.4059	.92704	21
Marketing Support	4.4861	.80948	24
Customer Service Support	4.6380	.86883	21
Other CRM Functions	4.7708	.64451	16
Total	4.5600	.82444	82

Levene's Test of Equality of Error Variances(a)

Dependent Variable: The Quality of DW Customer Relationship Data for Problem Enactment

F	df1	df2	Sig.
1.246	3	78	.299

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+item_61

Tests of Between-Subjects Effects

Dependent Variable: The Quality of DW Customer Relationship Data for Problem Enactment

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1.469[a]	3	.490	.713	.547	.027
Intercept	1679.615	1	1679.615	2444.816	.000	.969
item_61	1.469	3	.490	.713	.547	.027
Error	53.587	78	.687			
Total	1760.160	82				
Corrected Total	55.055	81				

a. R Squared = .027 (Adjusted R Squared = -.011)

Descriptive Statistics

Dependent Variable: Customer Relationship Problem Sense-Making Enhancements

61 What is the primary	Mean	Std. Deviation	N
Sales Support	5.9173	1.33560	21
Marketing Support	6.5768	1.07644	24
Customer Service Support	6.5373	1.10542	21
Other CRM Functions	6.6494	1.25341	16
Total	6.4119	1.20367	82

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Customer Relationship Problem Sense-Making Enhancements

F	df1	df2	Sig.
.666	3	78	.575

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+item_61

Tests of Between-Subjects Effects

Dependent Variable: Customer Relationship Problem Sense-Making Enhancements

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	7.023[a]	3	2.341	1.655	.184	.060
Intercept	3307.356	1	3307.356	2338.169	.000	.968
item_61	7.023	3	2.341	1.655	.184	.060
Error	110.332	78	1.415			
Total	3488.624	82				
Corrected Total	117.354	81				

a. R Squared = .060 (Adjusted R Squared = .024)

Descriptive Statistics

Dependent Variable: Intensity of data warehouse scanning searcl

61 What is the primary	Mean	Std. Deviation	N
Sales Support	45.8255	23.13579	21
Marketing Support	42.3125	20.60963	24
Customer Service Support	40.6986	18.35505	21
Other CRM Functions	38.2709	18.87700	16
Total	42.0102	20.22274	82

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Intensity of data warehouse scanning search

F	df1	df2	Sig.
.392	3	78	.759

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+item_61

Tests of Between-Subjects Effects

Dependent Variable: Intensity of data warehouse scanning search

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	567.718[a]	3	189.239	.453	.716	.017
Intercept	140041.331	1	140041.331	335.501	.000	.811
item_61	567.718	3	189.239	.453	.716	.017
Error	32557.973	78	417.410			
Total	177844.242	82				
Corrected Total	33125.690	81				

a. R Squared = .017 (Adjusted R Squared = -.021)

Descriptive Statistics

Dependent Variable: Intensity of data warehouse representativeness search heuristic

61 What is the primary	Mean	Std. Deviation	N
Sales Support	21.0028	14.56700	21
Marketing Support	21.6205	11.38206	24
Customer Service Support	22.6769	9.10324	21
Other CRM Functions	24.6223	9.73958	16
Total	22.3186	11.35269	82

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Intensity of data warehouse representativeness search heuristic

F	df1	df2	Sig.
.960	3	78	.416

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+item_61

Tests of Between-Subjects Effects

Dependent Variable: Intensity of data warehouse representativeness search heuristic

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	135.663[a]	3	45.221	.342	.795	.013
Intercept	40551.000	1	40551.000	306.969	.000	.797
item_61	135.663	3	45.221	.342	.795	.013
Error	10303.902	78	132.101			
Total	51285.310	82				
Corrected Total	10439.565	81				

a. R Squared = .013 (Adjusted R Squared = -.025)

Descriptive Statistics

Dependent Variable: Intensity of data warehouse availability search heuristic

61 What is the primary	Mean	Std. Deviation	N
Sales Support	22.0015	10.06638	21
Marketing Support	20.4517	8.73397	24
Customer Service Support	25.6406	11.12421	21
Other CRM Functions	26.6118	13.60252	16
Total	23.3794	10.85918	82

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Intensity of data warehouse availability search heuristic

F	df1	df2	Sig.
1.702	3	78	.173

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+item_61

Tests of Between-Subjects Effects

Dependent Variable: Intensity of data warehouse availability search heuristic

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	520.140[a]	3	173.380	1.497	.222	.054
Intercept	44979.526	1	44979.526	388.462	.000	.833
item_61	520.140	3	173.380	1.497	.222	.054
Error	9031.523	78	115.789			
Total	54372.601	82				
Corrected Total	9551.663	81				

a. R Squared = .054 (Adjusted R Squared = .018)

Descriptive Statistics

Dependent Variable: Intensity of data warehouse anchoring and adjustment search heuristic

61 What is the primary	Mean	Std. Deviation	N
Sales Support	23.8516	14.84845	21
Marketing Support	24.8904	16.39144	24
Customer Service Support	18.9639	7.60500	21
Other CRM Functions	20.6986	8.96584	16
Total	22.2887	12.88027	82

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Intensity of data warehouse anchoring and adjustment search heuristic

F	df1	df2	Sig.
2.461	3	78	.069

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+item_61

Tests of Between-Subjects Effects

Dependent Variable: Intensity of data warehouse anchoring and adjustment search heuristic

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	486.336[a]	3	162.112	.976	.408	.036
Intercept	39193.407	1	39193.407	236.038	.000	.752
item_61	486.336	3	162.112	.976	.408	.036
Error	12951.668	78	166.047			
Total	54174.426	82				
Corrected Total	13438.004	81				

a. R Squared = .036 (Adjusted R Squared = -.001)

Descriptive Statistics

Dependent Variable: Intensity of data warehouse positivy search heuristic

61 What is the primary	Mean	Std. Deviation	N
Sales Support	19.1349	7.34683	21
Marketing Support	20.8279	9.12691	24
Customer Service Support	19.4115	9.04306	21
Other CRM Functions	17.5918	8.83294	16
Total	19.4001	8.53751	82

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Intensity of data warehouse positivy search heuristic

F	df1	df2	Sig.
.147	3	78	.931

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+item_61

Tests of Between-Subjects Effects

Dependent Variable: Intensity of data warehouse positivy search heuristic

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	102.729[a]	3	34.243	.460	.711	.017
Intercept	29707.269	1	29707.269	399.423	.000	.837
item_61	102.729	3	34.243	.460	.711	.017
Error	5801.284	78	74.375			
Total	36765.988	82				
Corrected Total	5904.013	81				

a. R Squared = .017 (Adjusted R Squared = -.020)

FACTOR: Number of supported CRM data warehouse functions (DWFUNC)

Descriptive Statistics

Dependent Variable: The Quality of DW Customer Relationship Data for Problem Enactment

60 DUMMY Which CRM	Mean	Std. Deviation	N
Just one CRM function	4.5944	1.03297	12
More than one CRM Function	4.6257	.57362	12
Total	4.6101	.81728	24

Levene's Test of Equality of Error Variances(a)

Dependent Variable: The Quality of DW Customer Relationship Data for Problem Enactment

F	df1	df2	Sig.
13.368	1	22	.001

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+dummy60

Tests of Between-Subjects Effects

Dependent Variable: The Quality of DW Customer Relationship Data for Problem Enactment

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.006[a]	1	.006	.008	.928	.000
Intercept	510.068	1	510.068	730.715	.000	.971
dummy60	.006	1	.006	.008	.928	.000
Error	15.357	22	.698			
Total	525.431	24				
Corrected Total	15.363	23				

a. R Squared = .000 (Adjusted R Squared = -.045)

Descriptive Statistics

Dependent Variable: Customer Relationship Problem Sense-Makin Enhancements

60 DUMMY Which CRM	Mean	Std. Deviation	N
Just one CRM function	6.5132	1.34599	12
More than one CRM Function	6.5000	.93385	12
Total	6.5066	1.13296	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Customer Relationship Problem Sense-Making Enhancements

F	df1	df2	Sig.
1.394	1	22	.250

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+dummy60

Tests of Between-Subjects Effects

Dependent Variable: Customer Relationship Problem Sense-Making Enhancements

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.001[a]	1	.001	.001	.978	.000
Intercept	1016.054	1	1016.054	757.184	.000	.972
dummy60	.001	1	.001	.001	.978	.000
Error	29.521	22	1.342			
Total	1045.576	24				
Corrected Total	29.523	23				

a. R Squared = .000 (Adjusted R Squared = -.045)

Descriptive Statistics

Dependent Variable: Intensity of data warehouse scanning search

60 DUMMY Which CRM	Mean	Std. Deviation	N
Just one CRM function	49.9862	22.00723	12
More than one CRM Function	38.9446	18.28955	12
Total	44.4654	20.57710	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Intensity of data warehouse scanning search

F	df1	df2	Sig.
.167	1	22	.687

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+dummy60

Tests of Between-Subjects Effects

Dependent Variable: Intensity of data warehouse scanning search

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	731.510[a]	1	731.510	1.787	.195	.075
Intercept	47452.150	1	47452.150	115.903	.000	.840
dummy60	731.510	1	731.510	1.787	.195	.075
Error	9007.086	22	409.413			
Total	57190.746	24				
Corrected Total	9738.596	23				

a. R Squared = .075 (Adjusted R Squared = .033)

Descriptive Statistics

Dependent Variable: Intensity of data warehouse representativenes: search heuristic

60 DUMMY Which CRM	Mean	Std. Deviation	N
Just one CRM function	22.3382	9.58852	12
More than one CRM Function	28.0882	11.87304	12
Total	25.2132	10.95519	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Intensity of data warehouse representativeness search heuristic

F	df1	df2	Sig.
.623	1	22	.438

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+dummy60

Tests of Between-Subjects Effects

Dependent Variable: Intensity of data warehouse representativeness search heuristic

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	198.375[a]	1	198.375	1.703	.205	.072
Intercept	15256.969	1	15256.969	131.012	.000	.856
dummy60	198.375	1	198.375	1.703	.205	.072
Error	2561.997	22	116.454			
Total	18017.341	24				
Corrected Total	2760.372	23				

a. R Squared = .072 (Adjusted R Squared = .030)

Descriptive Statistics

Dependent Variable: Intensity of data warehouse availability search heuristic

60 DUMMY Which CRM	Mean	Std. Deviation	N
Just one CRM function	25.5859	13.08482	12
More than one CRM Function	20.4401	8.62082	12
Total	23.0130	11.15059	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Intensity of data warehouse availability search heuristic

F	df1	df2	Sig.
3.813	1	22	.064

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+dummy60

Tests of Between-Subjects Effects

Dependent Variable: Intensity of data warehouse availability search heuristic

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	158.878[a]	1	158.878	1.294	.268	.056
Intercept	12710.344	1	12710.344	103.534	.000	.825
dummy60	158.878	1	158.878	1.294	.268	.056
Error	2700.840	22	122.765			
Total	15570.062	24				
Corrected Total	2859.718	23				

a. R Squared = .056 (Adjusted R Squared = .013)

Descriptive Statistics

Dependent Variable: Intensity of data warehouse anchoring and adjustment search heuristic

60 DUMMY Which CRM	Mean	Std. Deviation	N
Just one CRM function	17.1152	5.33375	12
More than one CRM Function	22.5527	11.10351	12
Total	19.8340	8.96007	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Intensity of data warehouse anchoring and adjustment search heuristic

F	df1	df2	Sig.
2.622	1	22	.120

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+dummy60

Tests of Between-Subjects Effects

Dependent Variable: Intensity of data warehouse anchoring and adjustment search heuristic

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	177.398[a]	1	177.398	2.338	.140	.096
Intercept	9441.267	1	9441.267	124.443	.000	.850
dummy60	177.398	1	177.398	2.338	.140	.096
Error	1669.106	22	75.868			
Total	11287.772	24				
Corrected Total	1846.504	23				

a. R Squared = .096 (Adjusted R Squared = .055)

Descriptive Statistics

Dependent Variable: Intensity of data warehouse positivy search heuristic

60 DUMMY Which CRM	Mean	Std. Deviation	N
Just one CRM function	20.3402	10.95460	12
More than one CRM Function	17.9443	7.54549	12
Total	19.1422	9.28007	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Intensity of data warehouse positivy search heuristic

F	df1	df2	Sig.
.462	1	22	.504

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+dummy60

Tests of Between-Subjects Effects

Dependent Variable: Intensity of data warehouse positivy search heuristic

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	34.440[a]	1	34.440	.389	.539	.017
Intercept	8794.218	1	8794.218	99.405	.000	.819
dummy60	34.440	1	34.440	.389	.539	.017
Error	1946.314	22	88.469			
Total	10774.972	24				
Corrected Total	1980.754	23				

a. R Squared = .017 (Adjusted R Squared = -.027)

APPENDIX M: RESEARCH FINDINGS VALIDATION SEMI-STRUCTURED INTERVIEWS

A few days in advance interviewees received an e-mail (see bellow) with the supporting material for the interview. This consisted on (i) a MS-PowerPoint file with the quantitative research findings explained in a narrative and simple way, and (ii) another MS-PowerPoint file with the constructs' definitions. The research definitions file included an introduction to the research and the structured questions (i.e. open questions were not included):

Hi <name>,

Here you have my suggestions:
- Print out the file DEFINITIONS and have the pages at hand. Do not invest time reading/preparing yourself. You do not need to do it.
- It would be good if you have your PC on and the file FINDINGS open. Again, you do not need to read/prepare nothing before my call.

I will guide you w. my explanations. Pls., it would be good if you confirm that you can open both files.

Thanks.

RA

INTERVIEW SCRIPT AND INSTRUCTIONS

On the agreed day and time I phoned the interviewees following the next SCRIPT and provided the following instructions:

Introduction (5 minutes)

- Greeting and causal conversation in order to create a relax and friendly atmosphere
- Appreciation for participating in this step of the research
- Explain the confirmatory and explanatory objectives
- Explain the interview process
- Encourage for open and challenging opinions including disagreements.
- Introduce the short questionnaire with the structured questions. This questionnaire was in one of the slides in the DEFINITIONS file.

Interview per each research finding (85 minutes total).

1. I explained the research finding in a colloquial way. The interviewee had it in his screen
2. I asked for confirmation of understanding and provided clarifications when needed
3. I requested the interviewee for answering the structured question associated to the research finding. The questions had the following format "regarding finding #x I ...." and a

response in a 7-point Likert format (1-Strongly disagree, 2-Moderately disagree, 3-Slightly disagree, 4-Neither agree or disagree, 5-Slightly agree, 6-Moderately agree, 7-Strongly agree). The interviewee was kindly ask for (i) writing his response in the print out of the questionnaire and then (ii) share with me his answer.

4. Once, the interviewee provided his answer I updated him of the answers collected of former interviewees. I captured the comments.

5. Then, I asked the following nine open questions:

O1: Regarding finding #1a, Which, if any, would be the business benefits that you would attribute to the deployment of this measure of the quality on data warehouse customer relationship data for problem enactment in organizational settings?

O2: Regarding finding #1b: What, do you think about the scores on the quality on data warehouse customer relationship data for problem enactment?

O3: Regarding finding #2a, Which, if any, would be the business benefits that you would attribute to the deployment of this measure of the customer relationship problem sense making enhancements in organizational settings?

O4: Regarding finding #2b: What, do you think about the scores on the customer relationship problem sense making enhancements?

O5: Regarding finding #3: What, do you think about the finding that FOCUS is the predominant information search mode orientation?

O6: Regarding finding #4: What, do you think about the finding that TEMPLATE is the predominant information search heuristic orientation?

O7: Regarding finding #5a: What, do you think about the finding that quality on data warehouse customer relationship data for problem enactment causes customer relationship problem sense making enhancements in organizational?

O8: Regarding finding #5b: What, do you think about the finding that with the higher the customer relationship problem sense making enhancements in organizational is then the higher quality on data warehouse customer relationship data for problem enactment is?

O9: Regarding finding #5a &#5b: What, do you think about the finding that quality on data warehouse customer relationship data for problem enactment scored 4.54 and customer relationship problem sense making enhancements scored 6.37?

O10: Regarding finding #6: What, do you think about this model? (e.g. does it help understanding this situation?)

O11: Regarding finding #9: What, do you think about the finding that CRM job function is not a moderator in the model?

O12 Regarding finding #10: What, do you think about the finding that the number of supported CRM data warehouse supported functions is not a moderator in the model?

APPENDIX N: ADDITIONAL ANOVA ANALYSES TESTING GROUPS DIFFERENCES ON QUALITY VARIABLES FACTORED BY DWFUNC

FACTOR: Number of supported CRM data warehouse functions (DWFUNC). Additional research after qualitative findings

Descriptive Statistics

Dependent Variable: Data Task Utility

60 DUMMY Which CRM	Mean	Std. Deviation	N
Just one CRM function	4.2833	2.03149	12
More than one CRM Function	4.1333	1.37598	12
Total	4.2083	1.69857	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Data Task Utility

F	df1	df2	Sig.
7.083	1	22	.014

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+dummy60

Tests of Between-Subjects Effects

Dependent Variable: Data Task Utility

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.135[a]	1	.135	.045	.834	.002
Intercept	425.042	1	425.042	141.203	.000	.865
dummy60	.135	1	.135	.045	.834	.002
Error	66.223	22	3.010			
Total	491.400	24				
Corrected Total	66.358	23				

a. R Squared = .002 (Adjusted R Squared = -.043)

Descriptive Statistics

Dependent Variable: Customer Insights

60 DUMMY Which CRM	Mean	Std. Deviation	N
Just one CRM function	3.9167	2.38207	12
More than one CRM Function	4.2500	1.63067	12
Total	4.0833	2.00362	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Customer Insights

F	df1	df2	Sig.
4.068	1	22	.056

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+dummy60

Tests of Between-Subjects Effects

Dependent Variable: Customer Insights

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.667[a]	1	.667	.160	.693	.007
Intercept	400.167	1	400.167	96.040	.000	.814
dummy60	.667	1	.667	.160	.693	.007
Error	91.667	22	4.167			
Total	492.500	24				
Corrected Total	92.333	23				

a. R Squared = .007 (Adjusted R Squared = -.038)

Descriptive Statistics

Dependent Variable: Competitor Insights

60 DUMMY Which CRM	Mean	Std. Deviation	N
Just one CRM function	4.5278	2.00231	12
More than one CRM Function	4.0556	1.75714	12
Total	4.2917	1.85804	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Competitor Insights

F	df1	df2	Sig.
.531	1	22	.474

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+dummy60

Tests of Between-Subjects Effects

Dependent Variable: Competitor Insights

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1.338[a]	1	1.338	.377	.545	.017
Intercept	442.042	1	442.042	124.575	.000	.850
dummy60	1.338	1	1.338	.377	.545	.017
Error	78.065	22	3.548			
Total	521.444	24				
Corrected Total	79.403	23				

a. R Squared = .017 (Adjusted R Squared = -.028)

Descriptive Statistics

Dependent Variable: Data Enacment Utility

60 DUMMY Which CRM	Mean	Std. Deviation	N
Just one CRM function	4.7500	.63747	12
More than one CRM Function	4.8719	.48891	12
Total	4.8110	.55906	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Data Enacment Utility

F	df1	df2	Sig.
.656	1	22	.427

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+dummy60

Tests of Between-Subjects Effects

Dependent Variable: Data Enacment Utility

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.089[a]	1	.089	.276	.604	.012
Intercept	555.486	1	555.486	1721.386	.000	.987
dummy60	.089	1	.089	.276	.604	.012
Error	7.099	22	.323			
Total	562.675	24				
Corrected Total	7.189	23				

a. R Squared = .012 (Adjusted R Squared = -.032)

Descriptive Statistics

Dependent Variable: Data Equivocality

d60random	Mean	Std. Deviation	N
Just one CRM function	3.0417	1.71170	12
More than one CRM Function	2.4583	.91598	12
Total	2.7500	1.37525	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Data Equivocality

F	df1	df2	Sig.
2.069	1	22	.164

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+d60random

Tests of Between-Subjects Effects

Dependent Variable: Data Equivocality

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	2.042[a]	1	2.042	1.083	.309	.047
Intercept	181.500	1	181.500	96.314	.000	.814
d60random	2.042	1	2.042	1.083	.309	.047
Error	41.458	22	1.884			
Total	225.000	24				
Corrected Total	43.500	23				

a. R Squared = .047 (Adjusted R Squared = .004)

Descriptive Statistics

Dependent Variable: Data Integration

d60random	Mean	Std. Deviation	N
Just one CRM function	4.2222	1.64122	12
More than one CRM Function	4.6111	1.17923	12
Total	4.4167	1.41165	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Data Integration

F	df1	df2	Sig.
1.501	1	22	.234

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+d60random

Tests of Between-Subjects Effects

Dependent Variable: Data Integration

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.907[a]	1	.907	.444	.512	.020
Intercept	468.167	1	468.167	229.259	.000	.912
d60random	.907	1	.907	.444	.512	.020
Error	44.926	22	2.042			
Total	514.000	24				
Corrected Total	45.833	23				

a. R Squared = .020 (Adjusted R Squared = -.025)

Descriptive Statistics

Dependent Variable: Data Richness

d60random	Mean	Std. Deviation	N
Just one CRM function	5.7500	1.02025	12
More than one CRM Function	5.9938	.77603	12
Total	5.8719	.89518	24

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Data Richness

F	df1	df2	Sig.
.407	1	22	.530

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+d60random

Tests of Between-Subjects Effects

Dependent Variable: Data Richness

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.357[a]	1	.357	.434	.517	.019
Intercept	827.502	1	827.502	1007.224	.000	.979
d60random	.357	1	.357	.434	.517	.019
Error	18.074	22	.822			
Total	845.934	24				
Corrected Total	18.431	23				

a. R Squared = .019 (Adjusted R Squared = -.025)

APPENDIX O: DEFINITIONS

AA: See intensity of the data warehouse anchoring and adjustment search heuristic.

AVBLE: See intensity of the data warehouse availability search heuristic.

Availability search heuristic: It refers to assessing the probability of a situation as a function of prior situations [Tversky and Kahneman 1972, Wright 1980] (e.g. a marketer considering a series of occurences of actual costs incurred in past editions of a campaign when estimating the cost for a new edition of such campaign). The search ends once recent, salient information about a relevant precedent is found. (e.g.[Chi and Fan 1997]).

Anchoring and adjustment search heuristic: It refers essentially to the trial and error method [Tversky and Kahneman 1974, Chi and Fan 1997] (e.g. a marketer setting the price of a product starting with a baseline price and making a number of impact analysis in several of the cost components). The search ends once the adjustments are not improving an implicit/explicit value function on the information found (e.g.[Chi and Fan 1997]).

Behavior: The way in which an individual acts or works [Oxford 1993].

Belief: Personal ontological posture about some element of reality (e.g. [Rowland 1995]) formulated as a predicate on a subject cognitively constructed by individuals (e.g. [Pajares 2002]). Beliefs and knowledge are different in several aspects [Abelson 1979] (e.g. a belief is concerned with the existence or nonexistence of certain conceptual entity).

Capability: Capacity for doing something [Oxford 1993]. According to [Oxford 1993] the terms ability, capability and competence refer to the same concept.

Cognitive fit: Match between the needed types of information of a developed cognitive structure for performing a task (e.g. customer relationship problem enactment) and the types of available information (e.g. [Vessey 1991, Agarwal *et al.* 1996, Goodhue and Thompson 1995]).

Cognitive processes: Mental processes (e.g. perception, learning, memory, reasoning) involved in the acquisition, organization and use of information [Bandura 1994].

Cognitive structure (also cognition): Mental organized representation of information (e.g.[Cyert and March 1963, March and Simon 1958, Simon 1955, Walsh 1995, Argyris and Schön 1978]). Cognitive structure refers to the manner in which individual's knowledge is organized (e.g. levels of abstraction, decomposition, causal and functional relationships) while knowledge refers to what information is available. Therefore, according to [Wang and Chan 1995] the two concepts complement each other.

COMP: See competitor insights.

Competitor data insights: Data that is relevant for deriving information on the strengths, weaknesses, capabilities and strategies of competitors (e.g. [Narver and Slater 1990]).

Competitor insights (COMP): Manager's perception on the extent to which the data warehouse customer relationship data is a source of competitor data insights (e.g. [Day 1994]).

Context: Set of functional processes and resources involved in an organization (e.g. [Witteloostuijn 1996]).

CRM: See customer relationship management.

CRM job function (JOBFUNC): Customer relationship management job function primarily performed by the informant. Four CRM job functions were studied: Sales support, marketing support, customer services support and other CRM function.

CUST: See customer insights.

Customer data insights: Data that is relevant for deriving information on customer profile and future customer behaviour (e.g. [Narver and Slater 1990, Day 1994]).

Customer insights (CUST): Manager's perception on the extent to which the data warehouse customer relationship data is a source of customer data insights (e.g. [Day 1994]).

Customer relationship commitment: Customer's enduring desire to continue a relationship with a firm accompanied by this customer's willingness to make efforts at maintaining it (e.g. [Morgan and Hunt 1994]).

Customer relationship management (CRM): Cross-functional business process that drives customer value by the creation and maintenance of business to customer and end user durable, close and mutually beneficial relationships (e.g. [Leigh and Marshall 2001]. CRM involves market-sensing activities and customer-linking activities [Day 1994].

Customer relationship problem declarative knowledge enhancement (DKNOWE): Manager's self-belief about the extent that knowledge on customer relationship problems has improved due to the data warehouse customer relationship data. Declarative knowledge is the set of stored situational cues and facts (e.g., types of customers and selling situations) which allows the manager to recognize and categorize customer relationship situations (e.g. [Porter and Inks 2000, Leidner *et al.* 1999, Campbell 1994]).

Customer relationship problem enactment (ENAC): Manager's perception on the extent to which the data warehouse is a source of applicable and helpful customer relationship data for enacting customer relationship problems (e.g. [Fedorowicz and Lee 1998, Bailey and Pearson 1983, Venkatesh and Davis 2000]).

Customer relationship problem enactment procedural knowledge enhancement (PKNOWE): Manager's self-belief about the extent that one's skills enacting customer relationship problems have improved due to the data warehouse customer relationship data. Procedural knowledge consists of routines, actions, strategies, or heuristics that apply to a task domain (e.g. [Porter and Inks 2000, Campbell 1994]).

Customer relationship problem integrative complexity enhancement (ICPLXE): Manager's self-belief about the extent that the integrative complexity of the cognitive

structures about customer relationship problems has improved due to data warehouse customer relationship data. Integrative complexity is the level of comprehensiveness - i.e. number of factors in the cognitive structure- and connectedness -i.e. links among the factors in the cognitive structure (e.g. [McFadzean 1996, Wang and Chan 1995, Stone 1994]).

Customer relationship problem enacting self-efficacy (SELFE): The strength in the manager's self-belief in one's capabilities to execute given types of performances enacting prospective customer relationship problematic situations (e.g. [Bandura 1997, 1986]).

Customer relationship problem enactment knowledge enhancement (KNOWENH): Manager's self-belief about the extent that one's skills enacting customer relationship problems have improved due to the data warehouse customer relationship data.

Customer relationship problem sense making enhancements (ENACENH): Manager's self-belief about the extent that the enactment of customer relationship problems has improved due to the data warehouse's customer relationship data (e.g. [Weick 2000, DeLone and McLean 1992, Weick 1993]). The literature on this construct supports the dimensionality of this construct.

Customer relationship problem sense making information search behaviour: Manager's perceptions about behaviour in customer relationship problem sense making information search (e.g.[Weick 2000]). Customer relationship problem sense making behavior is a latent construct represented in this research by the following two behaviors: Customer relationship problem sense making information search mode and customer relationship problem sense making information search heuristic.

Customer relationship problem sense making information search heuristics: Manager's information search behaviour characterized by the type of a simplifying routine. The information search heuristics considered in this research are intensity of the data warehouse availability search heuristic (AVBLE), intensity of the data warehouse representativeness search heuristic (REP), intensity of the data warehouse

anchoring and adjustment search heuristic (AA), and intensity of the data warehouse positivy search heuristic (POSI).

Customer relationship problem sense making information search heuristic orientation (HEUR): Behavioral pattern in terms of the exhibited information search heuristics REP, AVBLE, AA, and POSI. This research has found that TEMPLATE and TRIAL-and-ERROR are two HEUR patterns

Customer relationship problem sense making information search mode: Manager's information search behavior characterized by the type of inquiry. The information search modes considered in this research are intensity of data warehouse scanning search mode (SCAN) and intensity of data warehouse focused search mode (FOCUS) (e.g. [Huber 1991, Aguilar 1967]).

Customer relationship problem sense making information search mode orientation (MODE): Behavioral pattern in terms of the exhibited information search modes SCAN and FOCUS. This research has found that FOCUSMO and SCANMO are two MODE patterns.

Customer relationship problem sense making information search focused mode orientation (FOCUSMO): Manager's information search behaviour mode orientation with an emphasis on intensity of data warehouse focused search mode (FOCUS) over intensity of data warehouse scanning search mode (SEARCH).

Customer relationship problem sense making information search scanning mode orientation (SCANMO): Manager's information search behaviour mode orientation with an emphasis on intensity of data warehouse scanning search mode (SEARCH) over intensity of data warehouse focused search mode (FOCUS).

Customer relationship problem sense making information search template heuristic orientation (TEMPLATE): Manager's information search behaviour heuristic orientation with an emphasis on intensity of data warehouse representativeness (REP) and availability (AVBLE) search heuristics over intensity of data warehouse anchor and adjustment (AA) and positivy (POSI) search heuristics.

This orientation involves an information template either as the reference while searching representative information or as the goal of the search.

Customer relationship problem sense making information search trial-and-error heuristic orientation (TRIAL-and-ERROR): Manager's information search behaviour heuristic orientation with an emphasis on intensity of data warehouse anchor and adjustment (AA) and positivy (POSI) search heuristics over intensity of data warehouse representativeness (REP) and availability (AVBLE) search heuristics. This orientation follows the trial-and-error approach.

Customer relationship satisfaction: Customer's affective state resulting from an overall appraisal of his or her relationship with a firm (e.g. [Anderson and Narus 1990]).

Customer trust: Customer's confidence in a firm's reliability and integrity (e.g. [De Wulf *et al.* 2001]).

Data: Symbols obtained through an encoding process of the environment (e.g. [Burke 1989]). Such symbols have not yet been evaluated for their worth to an individual in a specific situation-within-context (e.g. [McDonough 1963]).

Data enactment utility (ENACUT): Manager's perception on the extent to which the DW customer relationship data has enactment utility for the specific confronted situation. In our case, the situation is enacting customer relationship problems.

Data equivocality (EQUI): Manager's perception on the extent to which the data warehouse customer relationship data favors more than one interpretation for the enactment of customer relationship problems (e.g. [Goodhue *et al.* 2000, Jarke *et al.* 1999, Weick 1979, Daft and Weick 1984, Daft and Lengel 1986, Swanson 1987]).

Data insights: It refers to the attribution of some data's cognitive utility (e.g. [Kaplan and Simon 1990]).

Data integration (INT): Manager's perception on the extent to which the data warehouse customer relationship data is normalized in terms of data definitions and

logical data structures (e.g. [Goodhue *et al.* 1992, Goodhue *et al.* 2000, Peltier *et al.* 1998, Codd 1972]).

Data richness (RICH): Manager's perception on the extent to which the data warehouse customer relationship data is a source of customer relationship information (e.g. [Daft and Lengel 1984]).

Data task utility (TASKUT): Manager's perception on the extent to which the DW customer relationship data has instrumental utility for the task at hand. In our case, the tasks are in the context of customer relationship management.

Data trustworthiness (TRUST): Manager's perception on the extent to which the data warehouse customer relationship data is regarded as true and credible evidence for the enactment of customer relationship problems (e.g. [Jarke *et al.* 1999]).

Data warehouse: Integrated, non-volatile, collection of unrelated or disparate subject-oriented data sources where each unit of data is relevant to some moment in time and atomic or/and highly summarized (e.g. [Inmon 1996, Marakas 1998, Kelly 1997]).

Declarative knowledge: It includes facts, instructions, examples and concepts. It is knowledge that we can consciously recall [Anderson 1993].

DKNOWE: See customer relationship problem declarative knowledge enhancement.

DWFUNC: See CRM data warehouse function.

ENAC: See customer relationship problem enactment.

ENACENH: See customer relationship problem sense making enhancements.

Enactment: Sense making cognitive process that allows the generation of information, plausible interpretations of a (problematic) situation, and actions to be realized (e.g.[Weick 2000]). From my literature review I concluded that terms like "Problem statement" and "problem focus" [Kuhlthau 1993] refer to the concept of enacted problem.

ENACUT: See data enactment utility.

EQUI: See data equivocality.

FOCUS: See intensity of data warehouse focused search mode.

Focused search mode: The reactive behaviour people exhibit when they are looking for information specific to a problem to be addressed or question to be answered [Huber 1991].

FOCUSMO: See customer relationship problem sense making information search focused orientation.

HEUR: See customer relationship problem sense making information search heuristic orientation.

Heuristic: Simplifying routines used by people in their information processing activities in order to filter information coping with their cognitive limitations [Simon 1976]. People rely on a limited number of heuristics coping with uncertainty [Choo 1997, Newell and Simon 1972]. Selection or rejection of information is influenced by the individual's preferred heuristics [Tversky and Kahneman 1974, Hogarth and Makridakis 1981, Schwenk 1984, Hogarth 1987].

ICPLXE: See customer relationship problem integrative complexity enhancement.

Information: Data with an imparted contextual meaning by an individual (e.g. [Burke 1989, Goia 1986, Feldman and March 1981]) through enactment.

INT: See data integration.

Integrative complexity: Level of comprehensiveness -i.e. number of factors in the cognitive structure- and connectedness -i.e. links among the factors in the cognitive structure- (e.g. [Sullivan and Weaver 2000, Wang and Chan 1995, Feist 1994]).

Intensity of data warehouse focused search mode (FOCUS): Manager's perceptions on the personal's amount of effort in focused search on the data warehouse making sense of customer relationship problems. Focused search is the

reactive and directed information search behavior mode people exhibit when they are looking for information specific to a problem to be addressed or question to be answered.

Intensity of the data warehouse anchoring and adjustment search heuristic (AA): Manager's perceptions on the personal's amount of effort using anchoring and adjustment heuristic searching the data warehouse in order to make sense of customer relationship problems. Anchoring and adjustment heuristic refers essentially to the trial and error method (e.g. a marketer setting the price of a product starting with a baseline price and making a number of impact analyses in several of the cost components). This heuristic implies a recursive process and each step involves a search for additional information and an adjustment of the previous assessment. The search ends once the adjustments are not improving an implicit/explicit value function on the information found.

Intensity of the data warehouse availability search heuristic (AVBLE): Manager's perceptions on the personal's amount of effort using availability heuristic searching the data warehouse in order to make sense of customer relationship problems. Availability heuristic refers to assessing the probability of a situation as a function of prior situations (e.g. a marketer considering a series of occurrences of actual costs incurred in past editions of a campaign when estimating the cost for a new edition of such campaign). This heuristic implies the search for easily accessible information about relevant precedents. The search ends once recent, salient information about a relevant precedent is found.

Intensity of the data warehouse positivy search heuristic (POSI): Manager's perceptions on the personal's amount of effort using positivy heuristic searching the data warehouse in order to make sense of customer relationship problems. Positivy heuristic refers to confirming the probability of a situation using the trial and error method (e.g. a marketer looking for issues in accounts that were predicted to have a high risk). This heuristic implies a search for information that is fundamentally consistent with existing beliefs, theories and cognition. The search ends once the information found confirms the probability of a situation.

Intensity of the data warehouse representativeness search heuristic (REP): Manager's perceptions on the personal's amount of effort using representativeness heuristic searching the data warehouse in order to make sense of customer relationship problems. Representativenes heuristic refers (i) To assessing the probability of a situation as a representative of a category (e.g. a 'price-lowering by a competitor' situation can have common information with an 'attempt action to gain market-share' pattern), or (ii) to making generalizations based on new information about a sample -i.e. the sample is representative of a large population (e.g. to assess the national market success of a new product line based on the data likelihood ratio of a test market). This heuristic implies that the search ends when there is a satisficing fit between information about a situation and information about a category.

Intensity of data warehouse scanning search mode (SCAN): Manager's perceptions on the personal's amount of effort scanning the data warehouse making sense of customer relationship problems. Scanning search is the proactive and exploratory information search behavior mode people exhibit when they browse through information without a particular problem to solve.

Interpretation: Framed information, which contains a symbolic label attributed to a (problematic) situation (e.g.[Smith 1995, Wai-yi 1998, Thomas *et al.* 1993, Dervin 1992, Bruner 1990, Belkin 1980, Belkin *et al.* 1982]) Examples are customer oriented (e.g. [Day and Nedungadi 1994]), drama (e.g.[Corey and Wilson 1994, Burke 1969]), negative-losses (e.g. [Dutton and Jackson 1987, Tversky and Kahneman 1981]), threats and opportunities (e.g. [Kotler 1988]).

JOBFUNC: See CRM job function.

KNOWENH: See customer relationship problem enactment knowledge enhancement.

Knowledge: Information believed by an individual as "justified truth" (e.g. [Nonaka 1994]) and "stored" in memory (i.e. can be retrieved) in a cognitive structure (e.g. [Lamberts and Shanks 1997]) through learning. For the purposes of this research we adhere to Nonaka's view in that we focus on the individual's belief about the justification of knowledge and not on its truthfulness (i.e. individual's knowledge

might be falsifiable and not scientifically generalizable). Knowledge refers to what information is available in memory (e.g. [Wang and Chan 1995]).

Learning: The sequence of information processing activities conducted for the the creation of knowledge [Day 1994]. Furthermore, some authors believe that behavioural change is required for learning [Fiol and Lyles 1985, Levitt and March 1988, Huber 1991, Argyris and Schön 1978] while others believe that new ways of thinking are enough. [De Geus 1988] Others stress open-minded approaches to problem solving [Senge 1992], which includes new ways of formulating problem statements.

MODE: See customer relationship problem sense making information search mode orientation.

Number of supported CRM data warehouse functions (DWFUNC): Customer relationship management functions supported by the data warehouse. Four CRM functions were studied: Sales support, marketing support, customer services support and other CRM function.

Perceived environmental uncertainty: The absence of information about activities and events in the environment [Brannick 1998].

PKNOWE: See customer relationship problem enactment procedural knowledge enhancement.

POSI: See intensity of the data warehouse positivy search heuristic.

Positivy search heuristic: It refers to confirming the probability of a situation. This heuristic implies a search for information that is fundamentally consistent with existing beliefs, theories and cognition. The search ends once the information found confirms the probability of a situation [Evans 1989].

Problem: Individual's perception of a variance, or a gap, between the present and some desired state of affairs (e.g. [Simon 1977, Smith 1990]).

Problem statement: The formulation in linguistic terms of the problem elements and its structure [Smith 1989].

Procedural knowledge: It is based on skills to perform a specific task [Anderson 1993].

QUALITY: See quality of data warehouse customer relationship data for problem enactment.

Quality of data warehouse customer relationship data for problem enactment (QUALITY): Manager's perception on the extent to which the information derived from the available customer relationship data in the data warehouse fits with the manager's customer relationship management cognitive needs for problem enactment (e.g. [Goodhue *et al.* 2000, Huang *et al.* 1998]). It is the reverse construct of uncertainty enacting customer relationship problems. That is, perfect quality means zero uncertainty and vice versa.

Rationality (in a problem enactment situation): The extent to which the sense making process involves the collection of information relevant to the problem, and the reliance upon analysis of this information in enacting it (e.g. [Simon 1978]).

Relationship: Cognitive structure with information about the self, another person and the social interaction between the self and this other person [Berscheid 1994].

Relationship closeness: Interdependence of the partner's behaviours, including their emotions and thoughts [Kelley *et al.* 1983].

REP: See intensity of the data warehouse representativeness search heuristic.

Representativeness search heuristic: It refers (i) to assessing the probability of a situation as a representative of a category [Kahneman and Tversky 1972, Wright 1980], or (ii) to making generalizations based on new information about a sample [Wright 1980] (e.g. managers may quickly categorize a customer as representative of a segment). This heuristic implies that the search ends when there is a satisficing fit between information about a situation and information about a category (e.g. [Chi and Fan 1997]).

RICH: See data richness.

SCAN: See intensity of data warehouse scanning search mode.

SCANMO: See customer relationship problem sense making information search scanning orientation.

Scanning search mode: The proactive behavior people exhibit when they browse through information without a particular problem to solve [Aguilar 1967].

Self-efficacy: Self-belief in one's capabilities to organize and execute the sources of action required to manage prospective situations [Bandura 1977].

SELFE: See customer relationship problem enacting self-efficacy.

Sense making: High order cognitive process intended to reduce equivocality, or multiple meanings, in the information [Weick 2000]. For example, to make sense of a customer relationship problem means that heedful interrelating connects sufficient individual knowledge with situational demands [Weick 1993]. From my literature review, I concluded that the terms understand, make sense and comprehend refer to the same concept.

Situation: Individual's perception of the condition of an aspect in the task environment (e.g. [Oxford 1993]).

Skills: Learned cognitive capabilities with an associated proficiency at performing a task (e.g. [Kanfer and Ackerman 1989]).

Task environment: The set of immediate stakeholders such as customers and competitors which the focal organization has to directly interact with [Witteloostuijn 1996].

TASKUT: See data task utility.

TEMPLATE: See customer relationship problem sense making information search template orientation.

TRIAL-and-ERROR: See customer relationship problem sense making information search trial-and-error orientation.

TRUST: See data trustworthiness.

Type A data quality: Ignorance on the quality level that it is needed for the situation-within-context.

Type A uncertainty: Not knowing what information is needed [Brannick 1998].

Type B data quality: Being aware of the quality level that is needed for the situation-within-context but being unable to achieve it.

Type B uncertainty: Knowing what information is needed but being unable to access it [Brannick 1998].