Assessing the Cognitive Fit of Hypertext-based Learning Aids for Advanced Learning in Complex and Ill-structured Domains.

Alejandro Ramírez

A Thesis

in

The Faculty

of

Commerce and Administration

Presented in Partial Fulfilment of the Requirements
for the Degree of Doctor of Philosophy at
Concordia University
Montreal, Quebec, Canada

April, 1998

0-612-40306-8

ABSTRACT

Assessing the Cognitive Fit of Hypertext-based Learning Aids for Advanced Learning in Complex and Ill-structured Domains

Alejandro Ramírez, Ph.D.
Concordia University, 1998.

Leidner and Jarvenpaa's (95) theoretical views on the use of Information Technology (IT) to enhance Management School Education indicate that Hypermedia and the Internet have a "primary match" with both the Constructivist and the Cognitive Information Processing theories of learning, mainly because these technologies allow learners to create their own knowledge structures. This situation makes them compatible with Schein's (92) and Zuboff's (88) "visions" of IT in organizations when these visions are extended to include the impact of IT on electronic classrooms.

Literature regarding the use of hypermedia for training and education has been inconsistent when it comes to assessing its benefits. Some researchers indicate that there are advantages, others indicate that there are none. A similar situation in the Graphs versus Tables literature was settled by Vessey's (91) Cognitive Fit Theory. Within this theory, Vessey (91) indicated that when the tool provided matches the task at hand, performance increases in both accuracy and speed. If hypertext is considered as a tool to help in problem-solving, it is possible to use the Cognitive Fit Theory as an evaluation construct.

We empirically investigated the cognitive fit of hypertext-based learning aids for advanced learning in complex and ill-structured domains. This research was motivated by the following factors: (1) existing results regarding the benefits of hypertext-based learning aids are inconclusive; (2) hypertext has the potential to enhance the usefulness and usability of learning and training tools; (3) it is often argued that hypertext is useful in facilitating knowledge dissemination; and (4) there is no general agreement in how to evaluate hypertext as a learning and training tool.

Under the proposition that problem-solving with cognitive fit results in increased problem-solving efficiency and effectiveness, a theoretical framework is needed to provide answers to the research questions: (1) do hypertext-based learning aids enhance the acquisition of advanced knowledge? (2) will the advantages (speed and accuracy) of using hypertext-based learning aids increase when we increase the complexity of the material?

Two learning aids, a hypertext-based and a computer-based linear version were compared in an experiment. The learning material was the *Lucas, Ginzberg and Schultz' (90) Information Systems Implementation Model*. Speed and accuracy of subjects' performance were assessed quantitatively and qualitatively using verbal protocol analysis. Data were collected using a set of validated tools that include a test, two cases, a questionnaire, a tracing device, and verbal protocols.

The data analysis shows, as expected, a significant difference in efficiency and effectiveness between the groups at the highest level of complexity tested, and no differences at the lowest level. Verbal protocols show that subjects using the hypertext module had a deeper understanding of the relationships among the variables of the model.

TESTIMONIAL

At the end of this endeavour I would mention that a doctoral programme is indeed a collective effort. I have been fortunate in travelling down this road accompanied by magnanimous and caring people. Thanks to all of you that in one way or another helped me either by fulfilling your duties, being there on time and available, or by going out of your way to answer my needs.

Special thanks to the members of my committee: Dr. Thomas Jerome Tomberlin, my supervisor, who gave me his time and undivided attention each time I needed to meet with him, in an extremely busy schedule as Associate Dean of Research and Ph.D. Programme at the Faculty of Commerce and Administration. I learned from him at every interaction. He is a mentor and a friend. I am eternally indebted.

Dr. Suzanne Rivard, external member of my committee. Since the first time we met, I was impressed by her friendliness and accessibility. Being one of the most recognized researchers in Management Information Systems in Canada, she has been a model and a source of invaluable information. I admire her, not only for her success as a professional but also because she is an extraordinary human being. Her guidance has had a memorable impact on my career.

Dr. Steven G. Shaw, my friend, co-chair of my committee and expert on Educational Technology, my minor area. I regret not having the opportunity to play soccer together, or stopping at his office more often. Besides his wisdom, his constant motivation, his pertinent comments, and his witty observations, were enough to kept me going in the pursuit of new explanations, new horizons, new goals. At times I wondered why I did not enroll in an Educational Technology programme, which under his guidance seemed to be an appealing alternative.

Dr. Danièle Thomassin-Singh, external examiner, who read my thesis with extra care and made accurate comments and extremely relevant questions. She provided a detailed and helpful critical commentary of my work. Dr. Richard Schmid, internal examiner, who in brief statements enriched the implications of my research and in concise questions showed me the value of being extra careful when it comes to see the limitations of our work. Dr. Danielle Morin, who chaired the reviewing committee with her well known enthusiasm. Thank you!

To the members of the Faculty of Decision Sciences and Management Information Systems, and especially to Dr. Meral Büyükkurt who helped me during the dry phase of data collection. To Dr. Mohamed Khalifa, for bringing Hypertext to my attention when no one else was talking about it in Montreal. Thanks also to Johanne Hicks, Sue Callaghan and Diana, who were always happy to lend a hand when the doors seemed closed everywhere else. Sue worked professionally in the transcription of what looked like senseless verbal protocols.

To all my colleagues, specially *Ghassan Salamoun* for his friendship, availability and guidance, *Liette Lapointe* for her friendliness and encouragement, *Dat-Dao Nguyen* for his constant conversation on excellence and achievement, as well as for great food-for-thought at lunch time in the company of *Prof. Mahesh Sharma, Anamitra Shome,* and *Mehdi Farashahi*.

I want to dedicate this thesis to Carla Berend, my wife, for her love and support. Besides being the force that kept me going, she gave me three fantastic reasons to keep working every day: Daniel, Megan and Ian, our children. I knew before joining the programme that I needed a family to be able to pursue a doctorate. Now, I have one and the other. Daniel has been a source of wonder, Megan has shown me that caring is innate, and Ian has become a synonym of sweetness. I love you!

Juan Pablo, who has come to share with me his light, his dreams and his successes. It has been wonderful to be able to harmonize all aspects of my life just before closing this chapter.

I know that my parents planted a seed of success in me before they had to leave me in an ever changing world. Without them I would have never had a chance, but they said yes to life. I am deeply grateful and forever thankful. My sisters Inés, Ernestina, Martha Lourdes and María Dolores and my brother Mario for their love, understanding, and confidence. They always let me know they believed in me. This programme has kept us separated physically, but has allowed us to be closer emotionally.

Carla's parents, Stephen and Peggy Berend, who have been eager participants in this adventure and have been equally supportive and caring. As well as my brother-in-law, Michael and his family, Anne, Elizabeth and Robert.

Love to all my nephews, nieces, and their families: Saraberta, Analídise, Rodolfo Rafael, Iliana Ernestina, Claudia Patricia, Mario Alejandro, Omar, Efraín, Ana Lucía, and Isaac. Now they know that if there is a will, there is a way to achieve your dreams.

Finally, I would like to mention my gratitude to all my friends, and their respective families, that in one way or another kept me going, especially Carlos Briones, Jaime Bonilla, Alberto Rodríguez, Anabel Macías, Javier Gracia, François Riendeau, Philippe Provost, Jorge Ángel Díaz, Salvador Coutiño, Francis Scarpaleggia, Brad Romoff, Saad Feliciani and Joe Italiano.

Table of Contents

List of Figures

List of Tables

CHAPTER 1. INTRODUCTION

The arrival of the Information Age has turned the attention of researchers towards using the computer for almost every imaginable task and measuring its impact in many domains. The domains of training and learning are no exception. Research in these domains has merged two vast bodies of literature into the fields of computer-based training (CBT) and computer-assisted learning (CAL). These growing fields have captured the attention of researchers in several fields.

Historically, instructional usage of computers for training and learning has evolved through different stages: In the 1960s and 1970s, a variety of computer-assisted instructional (CAI) programs, currently referred to as Programmed Instruction, was developed. Most of these programs, however, did not meet the expectations that educators had for the new medium. Learning and training via computers was usually based on inflexible presentations of didactic material. Early teaching programs were limited by the hardware then available, and by the only available guide for instructional development: Behavioural Theory.

During the 1980s, the concept of Integrated Learning Systems (ILS) was created. ILS are the integration of hardware, software, and curricula for instruction. These systems are accountability driven, providing continuous performance feedback (i.e., number of questions answered correctly, incorrectly, completely and/or incompletely). Thus, even though trainees or learners may be working on a variety of topics, skills, or tasks, the system gives the trainer or instructor a very detailed analysis of what work has been done at each station.

Researchers in Artificial Intelligence and Cognitive Science have recently developed additional types of training and educational software: intelligent tutoring systems (ITS) and cognition enhancers or computer-based learning aids. Among the emerging cognition enhancers, we can group those known as empowering environments, hypermedia, and micro worlds (Brown, 85). The concept underlying a cognition enhancer is that the complementary cognitive strengths of a person and an information technology can be used in partnership (Salomon, Perkins, and Globerson, 91). The cognitive attributes of human beings give them an advantage over computers at applying peripheral real-world knowledge to ill-structured problems. Among our intelligent activities as humans, there is the ability to store and retrieve vast amounts of information efficiently, to solve complex problems or reach decisions, and to connect our thoughts and ideas in nonlinear, associative ways. Central to this behaviour are our complex organizational skills and our ability to adapt or modify our reactions based on reason and the use of skills given the situation at hand. If we want to use computers as tools to support these activities, they have to possess at least a subset of these abilities.

Today's computers are capable of storing, organizing and retrieving large amounts of information. The critical point now is their capacity of providing flexible navigation through complex processes and program functions that lead to intended solutions. Computers have to provide several nonlinear pathways that offer the necessary level of depth or breadth needed to explore a system at will. For educational purposes this means that the control or total direction of the learning aid is determined by the conditions at hand and at the discretion of the user, and not always predetermined by the developer. A program aiming to help in the learning process should therefore provide a flexible and nonlinear way of exploring several factors. To achieve this type of behaviour, the learning aid should contain some type of hypermedia implementation or some other kind of navigational aids.

Obviously, computer-based training and learning systems are not the solution to every problem in training and education. In the computer-based training and learning literature we can find encouraging results and discouraging ones. Dozens of studies show that CBT can improve learning, attitudes, and job performance (Adams, 93; Eberts, 88; Eberts and Brock, 84; Francis, 92; Kearsley, 85; Nelson and Palumbo, 92; Price, 91; Radlinski and McKendree, 92). Computer-based training also has some drawbacks. It is expensive to develop (Eberts, 88), learners may dislike it (Dionne and Krull, 92; Gery, 87; Rushby, 87), and it is difficult to assess. Still, there may be some issues that are better served if they are implemented in a computer-based learning aid. The question is how do we identify those issues? Methodologically, there is no solution to that question unless we define a way to evaluate them.

1.1 Using Hypertext for Computer-based Learning

As mentioned previously, computer-based learning has evolved from programmed instruction to some more complex content-focussed learning environments; mainly due to an evolution of computer systems. With faster, more powerful and cheaper systems, the applications have multiplied. Today, learning environments come in all sizes and shapes. They engage learners in authentic context-sensitive learning tasks, support collaborative learning activities and socially-negotiated interpretations of domain knowledge, among other issues. Examples of these environments include anchored instruction, computer-supported intentional learning environments, and cognitive flexible hypertexts (Jonassen, 93).

Anchored instruction's main goal is to ground knowledge acquisition in engaging, real-world, video-based problem-solving situations (Cognition and Technology Group, 92). *Computer-supported intentional learning environments* provide an exploratory environment to address subject matter and/or content area issues. This environment allows this exploratory process to be based on individual, self-determined needs and interests, and supports knowledge construction that is directed by the student and monitored through a group-supported collaborative environment that uses higher-order questioning, constructive feedback, and management skills (Scardamalia and Bereiter, 91).

Cognitive Flexible Hypertexts are intended to engage learners in more meaningful, transfer-oriented, advanced knowledge acquisition. Cognitive Flexible Hypertexts are supported by the Cognitive Flexibility Theory (Spiro et al., 88). This theory is a conceptual model for instruction based upon cognitive learning theory and attempts to avoid oversimplifying instruction, provides multiple representations of content, emphasizes case-based instruction, focuses on context-dependent knowledge, and supports the natural complexity of the content domain in order to foster the development of advanced knowledge, particularly in ill-structured knowledge domains.

Only recently have there been serious efforts in investigating whether hypertext could support the learning process or not (Jonassen and Mandl, 90; Tergan, 97). Results have not been conclusive. One reason for that is that the level of complexity of the task at hand has not been addressed. We believe that by including the level of complexity as part of the experimental design, it can be demonstrated that hypertext systems will offer some clear advantages for learning.

The potential for information technology to revolutionize education does exist.

Part of the reason that information technology has yet to revolutionize daily classroom activities is its rapid evolution. It is very difficult to keep current in a field that is only a support for the primary activity, which is teaching. With so many options, it's hard to find the right fit between the available technology and the goals of learning. So it is in this context that we are discussing the benefits of a particular technology - hypertext - in a very specific domain, i.e., when the domain is complex and ill-structured.

1.2 Research Questions

While potential users of hypertext-based learning aids can be instructional designers, educational technologists, teachers, and trainers, information systems professionals also can benefit from this research to address end-user training issues for managerial tasks. In the context of this research, hypertext systems are viewed as tools for problem-solving, as a means for improving decision-making processes and outcomes.

The principal research question is the following: do hypertext-based learning aids enhance the acquisition of advanced knowledge? More specifically, will the measures of performance (effectiveness and efficiency) increase as a result of learning with a hypertext-based learning aid? Since we believe that these measures are affected by the level of complexity of the material being learned, the question then can be rephrased as: will the measures of performance (effectiveness and efficiency) increase as a result of using hypertext-based learning aids when the level of complexity of the material is increased?

1.3 Importance of the Research and Motivation

Since the early versions of hypertext authoring tools, it has been argued that hypertext will have an impact on learning. In a way, by merely linking documents it is assumed that there will be major educational effects, that hypertext systems will solve most, if not all, educational problems, and that once students have access to hypertext they will dramatically improve in some almost magical way. This has prompted, obviously, a strong reaction known as the "hype" of hypertext, or "hyperexpectations"[1] These expectations have also prompted a reaction that considers hypertext just another *buzz word* and suggesting that those *promises* cannot be fulfilled.

Over the past decade, since the first ACM Conference on hypertext, *Hypertext 87*, several experiments have been conducted with mixed results. Most of those experiments have dealt with simple tasks, comparing hypertext to linear text. It is not surprising that at this level there was no evidence of significant differences in learning. Only when researchers such as Spiro et al., (87, 88), Spiro and Jehng, (90), Jonassen, (93), Jonassen and Mandl, (90), Jacobson, (94), and the Cognition and Technology Group at Vanderbilt University, (92), stressed the importance of designing complex issues without oversimplifying them, did the results show some differences.

The lack of compatibility of results may be no more than the lack of an agreement regarding how to assess the benefits of hypertext-based learning aids. Most of the arguments have considered hypertext as an end not as a tool. If hypertext is considered as a

---

1 The term was created by Chandler (90) in a different context, but it can also be used as a collection of high expectations.

tool to help in problem-solving, or decision-making, it is possible to use the *Cognitive Fit Theory* as an evaluation construct (Vessey, 91).

This research responds to the need of finding a solid construct to evaluate the advantages (if any) of hypertext-based learning aids. It also adds to the accumulation of empirical evaluation of alternative learning and training tools. Its results will be useful for assessing the effects of hypertext on learning, problem-solving, and decision-making, and for the design of hypertext-based learning aids. We hope to contribute to the design and evaluation of learning and training tools.

In summary, this research is motivated by the following factors: (1) existing results regarding the benefits of hypertext-based learning aids are inconclusive; (2) hypertext has the potential to enhance the usefulness and usability of learning and training tools; (3) it is often argued that hypertext is useful in facilitating knowledge dissemination; and (4) there is no agreement in how the advantages of hypertext as a learning and training tool should be assessed.

1.4 Organization of this dissertation

This dissertation is organized as follows: Chapter 2 reviews relevant prior research of the Cognitive Fit Paradigm, the use of information technology in post secondary education, Constructivism, Cognitive Flexibility Theory, and knowledge domains that are complex and ill-structured. This chapter also looks at the fit between information technology and learning by looking at specific visions of information technology that may be held by teachers. Chapter 2 ends with a review of the literature regarding the fit between

Constructivism and Hypermedia.

Chapter 3 presents a theoretical framework for evaluating the advantages of hypertext-based tools for learning and training under the Cognitive Fit Theory. Some specific issues regarding hypertext and learning are also presented in this chapter.

Chapter 4 develops the research model for guiding the proposed investigation, presenting the hypotheses based on the research model. Chapter 5 describes in detail the proposed research method, the task domain, the independent and dependent variables, the experimental design and procedures.

Research findings are presented in Chapters 6 and 7. The former reports on quantitative data analyses, the latter describes results of the verbal protocol analysis, and its qualitative interpretation.

Finally, Chapter 8 concludes the dissertation by discussing the major contributions, implications and limitations of this research. It also comments on directions for future research.

CHAPTER 2. REVIEW OF PRIOR RESEARCH

2.1 The Cognitive Fit Paradigm

Vessey (91) developed the notion that complexity in the task environment will be effectively reduced when the problem-solving aids support the task strategies required to perform that task. She termed that notion: *Cognitive Fit*. Her model was based on a general model of problem-solving depicted in Figure 1. The model views problem-solving as an outcome of the relationship between problem representation and the problem-solving task. Processes in her model are represented by the flows and arrows linking pairs of elements in the model. Also, in this model, she uses the mental representation as a problem representation in human working memory.

The Cognitive Fit Paradigm provides a theoretical background to research on information processing theory (Vessey and Galletta, 91). Newell and Simon (72) stated within their Information Processing Theory that human problem-solvers will strive to reduce their effort while solving a problem, since they are limited information processors. In order to facilitate the problem-solving process that human problem-solvers use in completing the task, the processing effort must be reduced. If the tool is matched to the task, the processing effort will be reduced. This is precisely the paradigm of Cognitive Fit.

The notion of Cognitive Fit presented in (Vessey, 91) applies to simple tasks. She was encouraged by the graphs versus table controversy in the MIS literature and wanted to settle the inconsistencies of the results. Some results suggested that subjects using graphs performed better than subjects using tables, other studies indicated exactly the opposite, while

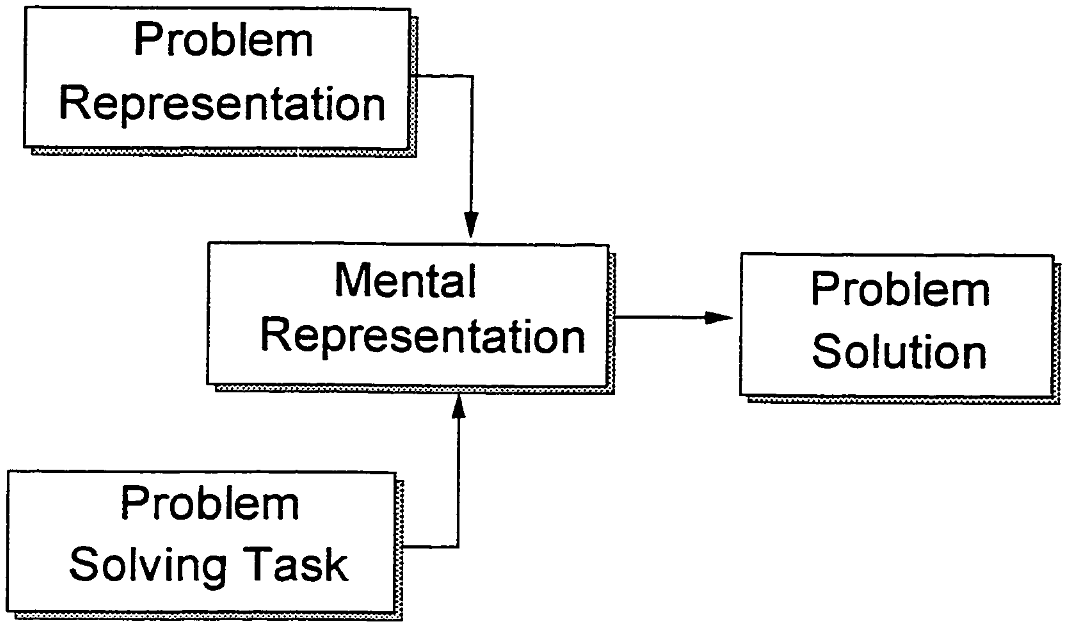


Figure 1 General Problem-Solving Model (Vessey, 91).

still other studies reported no differences. She gathered the results of all those studies and used them as her source of data. Using the cognitive fit paradigm, she proposed that "*problem-solving with cognitive fit results in increased speed and accuracy of performance.*"

In the discussion of findings, Vessey (91) claims that matching the problem representation to the type of task to be solved results in improved decision-making performance. The cognitive fit paradigm was successful in explaining the results of the graphs versus tables experiments, involving both information acquisition and information evaluation tasks. It was also successful in explaining most of the contradictory findings reported in the literature. Analysing her findings, it is possible to foresee two ways of examining more complex decision-making tasks: First, extend the notion of fit to more complex decision-making environments. Second, use the notion of elementary information and perceptual processes to characterize the diverse processing strategies involved when a number of possible sub-tasks are used to solve a problem. Accordingly, the possibility arises to use such a construct to assess the "match" (Venkatraman, 89) between any given task and the tool (or set of tools) used to perform that particular task. The assessment of the "match" is a way to validate the use of that tool (or set of tools) in that particular context.

Cognitive Fit is a cost-benefit characteristic that suggests that for most effective and efficient problem-solving to occur, the problem representation and any tools or aids employed should all support the strategies (methods or processes) required to perform the task (Vessey, 91). This means that the problem representation a problem-solver uses must be considered in the context of the task to be solved. Also, designers should concentrate on determining the characteristics of the tasks that the problem-solvers must address. Once these characteristics have been determined, they should be supported with the appropriate tools. For example, Vessey and Galletta (91) found that graphs are more appropriate tools when the task involves spatial problem-solving elements while tables are better suited if the task involves symbolic problem-solving elements. In other words, despite individual preferences, some tasks demand specific tools; If those tools are available, the task's goals will certainly be attained. This does not mean that there is only one way to obtain the task's goals. What it implies is that when cognitive fit exists, it confers advantages in performance.

Vessey and Galletta (91) investigated the effects of the basic paradigm of cognitive fit and extensions of the paradigm in a laboratory experiment that examined the nature of subjects' mental representations as well as problem-solving performance. In their experiment, they used 128 paid MBA students in two identical replications: an individual session and a class session. The individual session included an introduction to the experimental procedures, two practice exercises, five sets of experimental tasks, and a questionnaire to be completed. The class session included participant responses to standard questions involving either spatial or symbolic abilities.

The basic model of cognitive fit views problem-solving as the outcome of the relationship between the problem (external) representation and the problem-solving task. Vessey and Galletta (91) extended this notion of cognitive fit to include, in a similar way as problem representation, other problem-solving elements that potentially may have an effect on task solution. (Figure 2).

In the first part of the experiment Vessey and Galletta (91) tested whether more effective and efficient problem-solving results when the problem representation matches the task to be accomplished. In the case of spatial tasks, the proposition, graphs are more suitable than tables, was supported for time (it required less time to finish the task) but not for accuracy (solutions were faster but not necessarily correct). In the case of symbolic tasks, the proposition, tables are more suitable than graphs, was supported both for time and accuracy.

For the extended model, Vessey and Galletta (91) found that performance increased significantly with skills that supported either the task or both the problem representation and the task. However, there were no performance effects when the problem-solving skills supported the problem representation alone. Based on those results, they suggested a more detailed investigation into the characteristics of the type of tasks supported by the paradigm. The current research on cognitive fit applies to information acquisition and

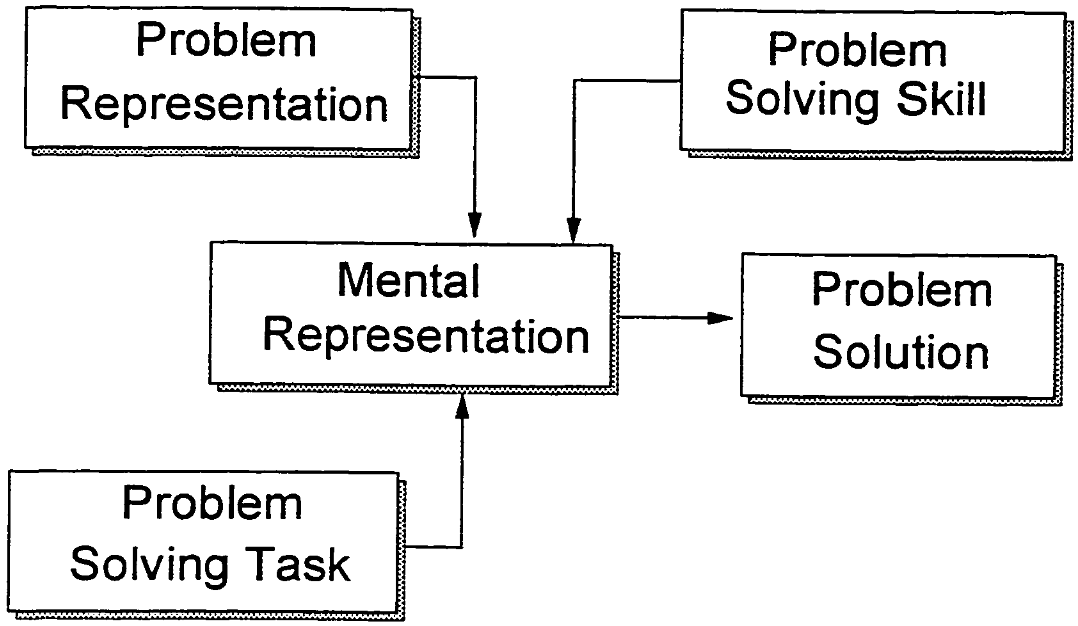


Figure 2 Extended Problem-Solving Model (Vessey and Galletta, 91).

simple information evaluation tasks. So they suggest as well that the research should be extended to encompass more complex problem-solving tasks.

One immediate implication for systems design within the Cognitive Fit Paradigm is that designers can examine the nature of the task to be performed and consequently support the task by providing the problem solver with the problem representation that matches the task (Vessey, 91). The management literature has long investigated the notion of fit (Joyce, Slocum and Von Glinov, 82; Venkatraman and Camillus, 84; Drazin and Van de Ven, 85; Alexander and Randolph, 85; Van de Ven and Drazin, 85; Venkatraman, 89). This investigation is useful since the notion of fit has been broadened. In this ample context, the notion of Cognitive Fit also benefits.

Cognitive Fit, then, can be used as a construct to measure the match of several task-tools pairs in different domains and/or contexts. It will provide enough confidence to researchers, designers, or practitioners to assign a particular tool in order to perform a given task. Unfortunately the literature on Cognitive Fit is limited. This research is offered as a continuation of that work.

2.2 Using Information Technology in Post Secondary Education

Even though information technology is not foreign to educators working at the post secondary level, its use has not directly changed the way they teach. Those few exceptions that have used information technology directly for teaching purposes have been developed by motivated educators who believe that IT will assist them in their teaching duties. Most of these examples are used without any evidence that they are actually having a favourable impact on the learning process.

Even today, when a large number of businesses are praising the benefits of using information technology in their activities, in schools and universities, lecturing still is the most common way to teach. The tools used are chalk, blackboards, overhead projectors, and paper-based handouts, even nowadays when computers are pervasive in the schools and universities. Virtually every school and university in the United States and Canada has microcomputers. Their use is still mainly to support teaching activities, not for teaching in itself. Slowly, the teacher's role in the teaching and learning process will change as new technologies are introduced into the classroom. There is an increasing pressure for universities to consider the adoption of information technology while, at the same time, concern is expressed for the impact of these technologies on individuals in society at large. So, this trend is a double-edged sword. Pleas for the use of information technology in schools and universities are increasing in frequency. At the same time, there are cries of concern over the impact of these technologies in almost every social activity.

2.2.1 Information Technology, Learning and Education

The process of learning, i.e., acquiring new knowledge and/or skills, has been under study since the beginning of civilization. How people learn is a question much older than organized Psychology, Biology, or Cognitive Science. Nonetheless there is no unique theory of learning. Learning, recognized by many as a key issue, is still eluding us. Recently, the literature on learning out of school, or non academic learning in school has grown substantially. The work of Willis (81), Scribner (84), Lave (88), Eckert (89), and Fosnot (89) challenges many assumptions about learning. Cognitive Science, Artificial Intelligence, and Intelligent Tutoring Systems have united efforts to build a practical theory of learning. Some accepted views are presented here.

Learning and Philosophy. Epistemology, one sub-discipline of Philosophy, is concerned with the nature of human knowledge. Epistemologists want to know what knowledge is, how to distinguish it from mere opinion and falsehood, how it is acquired, and how it is used. Obviously, learning is part of Epistemology. Many fields have inherited or drawn from epistemology a set of directions regarding knowledge and its acquisition, i.e., learning.

Learning and Biology. Biologists talk about changes to the nervous system of the organism, structures that make certain behaviours possible due to a particular history of interactions, e.g., the structures are ontogenic and the behaviour is described as learned. The criterion used to determine whether someone has knowledge is actually seeking an effective action in the domain of interest (Maturana and Varela, 87).

Learning and Behaviour. For some researchers the study of learning should be restricted to what is observable: One can see behaviour, one can see the environment, and one can see that behaviour adjusts itself to the environment. The study of learning might then be the study of *how* behaviour adjusts itself to the environment. This view that learning can be studied entirely in terms of environment and behaviour, or stimulus and response, is known as behaviourism. Within this view, learning is any relatively permanent change in a response pattern that results from experience or practice. After learning, a learner's behavioural repertoire changes. These changes often last a lifetime. Behaviourists neglect the claims made by biologists because if learning is determined by behaviour, then the internal history of how physiology underlies learning is unimportant.

Learning and Cognition. In an early stage, cognitive scientists compared the human mind to a computer, accepting input through perception, storing it in memory, processing it in thought, and acting on it (output) in making decisions. Later, cognitive theorists, i.e., researchers studying how people think, proposed that some learning takes place as the mind creates mental maps or charts to keep track of experience or stimuli. Within cognitive science there have emerged other theories, such as Anderson's (83) Adaptive Control of Thought (ACT) theory and Bandura's (77) Social Learning theory, among others.

Probably one of the most important contributions to learning within this approach, comes from Anderson's ACT theory, i.e., the bifurcation of knowledge into declarative and procedural. This distinction implies that there are differences in the way declarative versus procedural knowledge is acquired. The learner acquires declarative knowledge by adding a cognitive unit into short-term memory (STM) and then transferring it to long-term memory (LTM). Procedural learning is acquired by executing a skill; by doing. The implication is that procedural learning is a much more gradual process than declarative learning (Anderson, 83).

Stanford's psychologist Albert Bandura and others brought several elements of conditioning and cognitive theories together in developing what Bandura (77) calls Social Learning Theory. The basic premise is that behaviour results from a continuous interaction between several key determinants, also known as significant factors. These include imitation, observational learning, vicarious learning, and symbolic learning. But the learner also plays a prominent role in cognitively selecting, organizing, and transforming stimuli from the environment. This developmental view postulates one or several autonomous, inner-directed, developmental processes that guide learning at each stage of development.

For many researchers knowledge is a process of social construction (Wittgenstein, 53; Bloor, 83; Winograd and Flores, 86; Maturana and Varela, 87; Lave, 88). The concept of information transfer as a metaphor for knowledge acquisition is under severe questioning. For them, learning is much more an evolutionary, sense-making, experiential process of development than the simple acquisition or transfer of information (Maturana and Varela, 87; Varela, 89). Then, how can we assess student learning to the end that after the student has learned, his or her behaviour has changed? How can we know that? This is possible by observing the ability to do something, a new or improved mental ability, and/or a change in attitude.

As Norman (81) puts it: "*We spend much of our lifetime learning: in a sense we learn from every thing we do. If learning is not yet understood, it is because there is more to it than the simple accumulation of knowledge.*" Accumulation is only one aspect of learning, there are several other aspects of it that we have to understand and explain. One aspect this dissertation particularly deals with is that of the restructuring of acquired knowledge; how the very basis of understanding of some piece of information changes as a result of new concepts or new experiences.

2.2.2 *Constructivism*

In brief, Constructivism is a theory of knowledge derived from the philosophical proposition that reality is created, or constructed, by the individual. Constructivism claims that reality is more in the mind of the knower; that the knower constructs a reality, or at least interprets it, based on her own perception. While the emphasis in objectivism is on the object of our knowledge, constructivism is concerned with how we construct knowledge. Knowledge construction varies with experiences, mental structures, and beliefs that one uses to interpret objects and events.

Radical constructivists believe that there is no real world, no objective reality independent of human mental activity. A less radical approach holds that the mind is instrumental and essential in interpreting events, objects, and perspectives in the real world, and that those interpretations comprise a knowledge base that is personal and individualistic. According to these less radical constructivists, thinking is grounded in perception of physical and social experiences, comprehended only by the mind. What the mind produces are mental models that explain to the knower what he or she has perceived. One important epistemological assumption of these theorists is that meaning varies with how the individual creates meaning from his or her experiences.

The constructivists' views can be summarized as: (1) There is a real world that we experience. (2) Meaning is imposed on the world by us, rather than existing in the world independently of us. (3) Meaning is seen as rooted in, and indexed by, experience (Brown et al., 89). (4) Constructivists emphasize "situating" cognitive experiences in authentic activities, a good example of this view is the concept of "Cognitive Apprenticeship" developed by Collins et al., (89). (5) There is no ultimate, shared reality, but rather a reality that is the outcome of constructive processes.

A major assumption of constructivism is that individuals learn better when they discover things themselves and when they control the pace of learning. The purpose of instruction moves away from knowledge dissemination towards knowledge creation. The process of creation is more important than the knowledge created, since it can be redesigned later, when new connections or structures are developed.

Following the interpretation of Spiro et al. (88) we see constructivism in learning and education as a new understanding, constructed when various sources of knowledge are put together. With such an interpretation, they developed the Cognitive Flexibility Theory, which adds a new element of constructive processing to those already accepted within Constructivism, an element concerned mainly with the flexible use of pre-existing knowledge (i.e., new understandings are constructed by using prior knowledge to go beyond the information given; pre-existing knowledge is itself constructed on a case-by-case basis).

2.2.3 *Cognitive Flexibility Theory*

Cognitive Flexibility Theory is a conceptual model for instruction based upon cognitive learning theory and attempts to avoid the oversimplification of instructions, provides multiple representations of content, emphasizes case-based instruction, focuses on context dependent knowledge, and supports the natural complexity of the content domain to foster the development of advanced knowledge, particularly in ill-structured knowledge domains (Spiro et al. 88). Cognitive Flexibility Hypertexts are designed to engage learners in more meaningful, transfer-oriented, advanced knowledge acquisition.

The central claim of Cognitive Flexibility Theory is that revisiting the same material, at different times, in rearranged contexts, for different purposes, and from different conceptual perspectives is essential to attain the goals of advanced knowledge acquisition. Content must be covered more than once for full understanding to a point that associations among elements will become apparent. If some of these associations are not grasped on a first exploration, they may be noticed on a second or third try.

If we take into account all the previously mentioned factors, it is simple to see that the main metaphor employed in the instructional model derived within Cognitive Flexibility Theory is that of the crisscross landscape (Spiro et al. 88). This metaphor suggests a nonlinear and multidimensional traversal of complex subject matter, returning to the same place in the conceptual landscape on different occasions, coming from different directions. The instructional model suggests that hypertext, with its nonlinear properties, is a viable technology for building this crisscross landscape.

2.2.4 Complex and Ill-structured Knowledge Domains

An ill-structured domain is determined by its conceptual complexity and its across-case irregularities. In other words, each case or example of knowledge application typically involves multiple conceptual structures (multiple schemas, perspectives, organizational principles, etc.), each of which is individually complex (including both concept and case complexity). The pattern of interaction varies substantially across cases of the same type (each case looks like a 'new' one).

Examples of ill-structured domains include medicine, history, art appreciation, literary interpretation, system's modelling, and systems design. Furthermore, Spiro et al. (92) argue that *all domains which involve the application of knowledge to unconstrained, naturally occurring situations are substantially ill-structured.* They give as an example, the case of engineering, which employs basic principles of Physics (most of them well-structured in the abstract) to "messy" real-world cases. The nature of each case in engineering is so complex and differs so much from other cases that it is difficult to categorize them under any single principle.

There are only a few studies investigating the process of learning complex topics. Most of the literature deals with learning at an introductory level or with expertise. Spiro et al. (88) discuss some special characteristics of advanced learning of complex conceptual material. They provide a rationale about an intermediate stage, between introductory learning and expertise, in a knowledge domain where the learning goals must change, where at some point students must "get it right." At this stage, "the learner must attain a deeper understanding of content material, reason with it, and apply it flexibly in diverse contexts." One characteristic of advanced knowledge is that its conceptual complexity increases as it becomes less structured.

Therefore, learning to carry out a procedure, use a concept, understand a system, in this context becomes onerous. Simple examples, memorization and simplification are no longer useful. Knowledge here is intertwined and dependent, has significant context-dependent variations, and requires the ability to respond with flexibility to complex situations.

2.3 The Fit between Information Technology and Learning

In a theoretical review of the use of information technology to enhance management school education, Leidner and Jarvenpaa (95) looked at the pedagogical assumptions underlying the design of information technology for educational purposes. They argue that in this context the use of IT has followed almost exclusively the automation paradigm. In an effort to stimulate innovative applications and rigorous examinations of the effectiveness of these applications, they discussed five models of learning in the context of management education: objectivism, constructivism, collaborationism, cognitive information processing, and socioculturalism.

After a brief description of each one of those models of learning, they are mapped onto specific sets of technologies that enable instructors to informate up and down and ultimately transform the educational environment and process. We are using this theoretical foundation in our work, since we are interested in the use of information technology - hypertext - to improve the learning process within the constructivist model of learning.

2.3.1 *Educational assumptions about Information Technology*

Educators using information technology will adhere, consciously or unconsciously, to some educational assumptions about it. Because their understanding of the educational process can be classified as behavioural or cognitive, they will see the benefits of information technology in education differently. Leidner and Jarvenpaa (95) looked at five learning models: objectivism, constructivism, collaborationism, cognitive information processing, and socioculturalism. We are interested exclusively in constructivism, but for the sake of clarity, will discuss briefly the other four models, since we have already covered constructivism in section 2.2.2.

The objectivist model of learning is based on Skinner's (53, 69, 71) stimulus-response theory. The goals of learning here are to understand reality and modify our behaviour accordingly. Reality is seen as an objective entity waiting to be understood. Educators under this model believe that expertise is an accumulation of knowledge and we need to be able to transfer this knowledge to novices. Here, knowledge is also an objective entity that can be stored and transmitted. So the main concern of educators in this model is to find an efficient way to do so. They believe that instructors are the source of knowledge and they should be in control of the learning process.

The main difference between constructivism and objectivism is that reality is understood differently by either theory. Constructivism sees reality as a subjective experience. Knowledge is not transmitted, but created in a series of interactions by each learner. This model calls for a learner-centred approach. Instructors are facilitators of knowledge, but learners should be in control of the process. Individuals learn better when they are forced to discover things themselves rather than by being told or instructed.

For the educators that adhere to the collaborationist model of learning, learning emerges through shared understanding of more than one learner. In other words, learning does not exist in isolation. So, in order to generate knowledge, socialization is promoted, as well as group skills, communication, listening, and participation. Involvement is seen as critical for learning. In this context, instruction is communication-oriented, and the instructor is seen as a questioner and a discussion leader.

The Cognitive Information Processing model looks at learning as the creation and transfer of new knowledge into long-term memory. Its goal is to improve the cognitive processing abilities of learners, by improving recall and retention. Here, prior knowledge affects the level of instructional support needed, and educators believe that learners have limited selective attention. So they are concerned about instructional aspects of stimuli that can affect attention. They need some kind of feedback on student's learning in order to adapt to their needs.

Finally, those who adhere to socioculturalism see learning as subjective and individualistic, and as a means to empowerment, and emancipation. They believe that action-oriented, socially conscious learners will change society rather than merely accept or understand it. Then, instruction is always culturally value-laden, and embedded in a person's every day cultural or social context.

2.3.2 Specific Visions about Information Technology

It is easily understood that educators having different views of the education process will have different assumptions about how information technology can help them achieve their goals. In order to classify these assumptions, following Leidner and Jarvenpaa's (95) approach, we are going to look at different ways that educators look at information technology: As a means to automate, informate up, informate down, and transform. These "visions" are based on Schein's (92) organizational research in information technology.

Educators, whether they are explicitly aware of it or not, have a set of assumptions about information technology and a vision of what it can or cannot do for them. Many have a positive vision of information technology in general, but there is a great diversity in the nature and strength of that vision.

2.3.2.1 The Vision to Automate

When educators see information technology as a way of replacing expensive, unreliable human labour with sophisticated systems and other information technologies, they have a vision to automate. They see in these technologies a way to save money, to improve quality, and to make schools more effective. They acknowledge that teaching and learning cannot be automated, but they believe there are some aspects that can be better served using computers, particularly the delivery of information.

Within this vision, Leidner and Jarvenpaa (95) grouped the following technologies, since they are seen as prone to provide tools for manipulating and representing instructional material in a classroom: Instructors consoles with or without stand-alone student computers, Computer-Assisted Learning, and Distance Learning technologies.

2.3.2.2 The Vision to Informate Up

Zuboff (88) talks about the impact that information technology has in making previously concealed parts of a system's process more visible to people both higher up and lower down in the organization. She calls this process: to Informate. By adapting this term for education we can argue that, for some educators, information technology can become the ultimate educational control tool. They assume that by installing the right kind of information system they can monitor the educational process, pinpoint problems rapidly and prescribe remedial measures. Technologies that fit this purpose, identified by Leidner and Jarvenpaa (95) include key response pads[2] and any kind of E-mail system that allows communication between instructors and students.

---

2 Key response pads allow a large class of students to participate by responding to questions with a yes/no response or rating agreement to an issue on a 0 to 9 scale.

2.3.2.3 The Vision to Informate Down

Derived from the same principle as to informate up, the vision to informate down differs from the previous one mainly in what Zuboff (88) observed to be the consequence of introducing computers in the production processes and creating automated factories. The understanding of the production process, as a whole, is no longer in the hands of supervisors. This understanding is now distributed among the entire work force, shifting the power structure of the organization. In an educational context, this means that instructors now share the responsibilities of the learning process with learners, mainly by providing students greater access to information. Students now can critically analyse information and discuss issues between themselves and with the instructor. The technologies identified within this vision include Hypermedia, the Internet, Simulations, Virtual Reality, and Classrooms with synchronous communication devices either with or without group support (Leidner and Jarvenpaa, 95).

2.3.2.4 The Vision to Transform

According to Schein (92), few CEOs see information technology as the basis for a complete transformation of their organization and industry. They do see how these technologies can change organizations by altering the nature of the products, markets, and organizational structure, and how these in combination may alter organizational boundaries, inter-organizational relationships, and even the management process itself. In an educational context these issues will translate into a new understanding of the classroom, by making its physical boundaries obsolete; of teamwork, by having more effective interactions; of the learning process, by allowing it to become a continuous, time-independent process; and of the creation of knowledge, by enabling it to be a multilevel, multi-speed process. Technologies that Leidner and Jarvenpaa (95) identified within this vision are virtual learning spaces, known as Asynchronous Communication Across Distances (ACADs) either with or without group support.

Table 1, summarizes these visions. It also shows the convergence of Constructivism with Hypermedia, which is the area of interest of this dissertation. We can see that according to Leidner and Jarvenpaa (95) Hypermedia has a primary fit with both the Constructivist and Cognitive Information Processing learning models. This theoretical basis will be used to support our assessment of the cognitive fit of hypertext-based learning aids for advanced learning in complex and ill-structured knowledge domains.

In order to assess the impact of information technology on learning, Leidner and Jarvenpaa (95) evaluated each one of these visions against the five theories of learning regarding control of the pace and content of learning, and the purpose of instruction. They suggest that IS researchers, interested in this area, can benefit by using well-established variables from educational research rather than creating new ones. By examining well-defined learning outcome variables as dependent measures in studies of the impact of information technology on learning, their findings will be more comparable and easier to interpret.

	Objectivist	Constructivist	Collaborative	Cognitive IP	Sociocultural
The Vision to Automate					
Instructor Console	✔✔				
Instructor Console and Stand alone student Computers	✔✔	✔			
Computer-Assisted Learning	✔✔			✔	
Distance Learning	✔✔				✔
The Vision to Informate Up					
Key Response Pads	✔✔			✔	
Instructor-Student E-mail				✔	
The Vision to Informate Down					
Learning Networks		✔✔		✔	
Hypermedia/Internet		✔✔		✔✔	
Simulation/Virtual Reality		✔✔			
Synchronous Communication Classrooms (SCCs)			✔✔		✔
Groupware-supported SCCs			✔✔	✔✔	✔
The Vision to Transform					
Asynchronous Communication Across Distances (ACADs)			✔✔		✔
Groupware-supported ACADs			✔✔	✔✔	✔

✔✔ Represents the Primary match. **✔ Represents a secondary match.**

Table 1: Information Technology Fit with the Theories of Learning (Leidner and Jarvenpaa, 95)

2.3.3 The Fit between Constructivism and Hypermedia

Hypermedia[3] places much of the control of the content and pace of learning in the hands of students. The purpose of study moves away from knowledge

---

[3] Hypermedia and hypertext are being used interchangeably throughout this dissertation even though hypermedia includes audio, video, and graphics in addition to text.

dissemination to knowledge creation. Much of the knowledge has already been created, i.e., is explicit, but the instructor is no longer the primary creator of it. Then its structure becomes very important and students become part of the knowledge creation process, while the instructor becomes a mediator rather than an adviser of the learning process.

From Table 1, we can conclude that hypermedia is more properly used in constructivist or cognitive information processing environments, with an emphasis on conceptual learning and higher-order thinking. That was also noted previously by Jonassen (93), Duffy and Jonassen (92), and Spiro and Jehng (90). Since hypermedia is a critical part of this research, the rest of this chapter is dedicated exclusively to hypermedia and research dealing with hypermedia.

2.3.3.1 Hypertext - Hypermedia

Hypertext, a computerized form of information, is seen as nonlinear prose, interactive print, or dynamic text. It is usually defined as the nonlinear viewing of information. "Nonlinear" means that information can be examined in any order by selecting the topic to see next. It provides a new way of accessing and organizing any type of information. Hypertext can make it easier and faster to find things and absorb ideas.

Researchers and practitioners of hypertext claim that it will revolutionize information systems and cause designers of information to reexamine widely held beliefs and practices relating to publishing, text and screen design, information access and retrieval, user support, and documentation management. Hypertext's cognitive power comes from visually representing the hyper web on the screen to show available links and nodes or updated user paths. Hypertext provides a mental model for representing ideas by transforming the unobservable into objects that can be examined and inspected. Hypertext systems are being used extensively nowadays in many fields, from software engineering to collaborative problem-solving applications, as online documentation, information retrieval and help systems, as writing aids, and more recently as authoring tools for instruction and learning.

Results from preliminary studies have suggested that hypertext-based systems can be usefully applied in training and education and that it is worthwhile to generate further material for this medium (Locatis et al. 89; Hammond 89; Duffy and Knuth 90; Ess 91; Jones and Spiro 92; Michalak and Coney 93; Madigan et al. 94). Although all these research studies have proven the usefulness of hypertext technology for training and learning, still no one has demonstrated any appreciable advantage for hypertext. Like many new technologies brought to education, more claims have been made on behalf of hypertext than the experimental evidence has been able to support. Our position here is that hypertext is not the solution for every problem in education, but it is a solution for some problems; the sort of problems that by using hypertext's unique structure either reduce the complexity of the problem and/or help learners to understand them better. The challenge is twofold: to identify those problems, and to implement them successfully in hypertext.

2.3.3.2 Hypertext versus Linear text

A number of studies have been conducted comparing the effectiveness of hypertext and linear text. Some studies involved both the electronic and print media. Others, such as the current study, involve only one of the two media. In some of the studies the goal was to evaluate the effectiveness of hypertext as a search tool, i.e., a tool for locating specific information within a document. Typically in these studies subjects would receive a set of questions to be answered. There are only a few studies dealing specifically with the question of the educational effectiveness of hypertext as compared to linear text. These studies are reviewed in the next section.

2.3.3.3 Research Studies dealing with Hypertext

A few studies have been conducted in order to compare hypertext-based educational systems to linear text-based ones. After the first ACM (87) conference on Hypertext, held at The University of North Carolina - Chapel Hill, there was an explosion of interest about hypertext in general, and specifically on its educational implications. At Brown University, a group of researchers developed the Hypertext Editing Support System (HESS), the File Retrieval and Editing Support System (FRESS), and Intermedia, three hypermedia systems. These systems were used in a series of empirical studies (Catano, 89; Van Dam, 88; Yankelovich et al., 88). One of them, conducted with FRESS, involved a course on poetry. In that study, subjects were allowed to read, on-line, the hypertext material. The purpose was to compare their performance compared to other sections of Brown University's introductory poetry course. Results showed that subjects in the experimental group wrote three times as much for both analysis and informational discourse as did subjects in the control group (Catano, 89; Van Dam, 88), even though results were unreliable due to a very small experimental group (n=12).

Also, at Brown University, several experiments were conducted to evaluate the usefulness of hypermedia applications developed through Intermedia. Courses in English literature and plant cell Biology formed the instructional domains of the experiments. These courses included approximately eighty students. Yankelovich, et al., (88) reported a substantial increase in the critical thinking skills of subjects working with the hypermedia material.

Gordon, et al. (88) conducted a study with twenty-four subjects in order to evaluate the use of hypertext as an intra document text format. Subjects were asked to read two articles, one in linear text, the other in hypertext. A 2X2 mixed factor design was used, including two treatments (general interest vs. technical articles) and two factors (linear text vs. hypertext). Both factors were computer-based. After reading the articles, subjects were tested using three instruments: a free recall test, a question probe test, and a questionnaire. Findings indicated that linear text subjects were more successful remembering basic ideas for both types of articles, and showed a better assimilation of the macro-structure contained in the articles labelled as general interest. Subjects indicated a preference for linear text indicating that hypertext required more effort, while reading the material.

In a study to compare the performance of subjects exposed to hypertext material in an introductory MBA course in MIS, subjects were randomly assigned to two groups. Group one used the hypertext material to prepare for two non consecutive lectures, while group two used it for the lecture in between the lectures assigned to group one. That design allowed the authors (Khalifa and Ramírez, 92) to perform within-group as well as between-group comparisons. Each lecture was tested by means of an exam; the subject's mark in a given exam was used to measure their performance. Results were only conclusive in the between subjects analysis on the third exam. Each group obtained a higher average mark on the lectures where they used the hypertext material.

Table 2 contains a summary of the previous studies. As we can see, results regarding the benefits of hypertext-based systems are contradictory. One reason we can mention is that in those studies, they were addressing objectivist instead of constructivist learning. In that context the benefits of hypertext are not clear. Also, the learning material was quite simple and the lack of complexity made it equally suitable for linear and hypertext applications.

STUDY	RESULTS
Catano, 89; Van Dam, 88	Subjects in the Hypertext group wrote three times as much for both analysis and informational discourse as did subjects in the Linear text group.
Yankelovich, et al., 88	A substantial increase in the critical thinking skills of subjects working with the hypermedia material.
Gordon, et al., 88	Linear text subjects were more successful remembering basic ideas for both types of articles, and showed a better assimilation of the macro-structure contained in the articles labelled as general interest. Subjects indicated a preference for linear text indicating that hypertext required more effort, while reading the material.
Khalifa and Ramírez, 92	Results were only conclusive in the between subjects analysis on the third exam. Each group obtained a higher average mark on the lectures where they used the hypertext material.

Table 2: Summary of Research Studies comparing Hypertext versus Linear text.

CHAPTER 3. THEORETICAL PERSPECTIVES

In making an assessment regarding the Cognitive Fit of Hypertext-based Learning Aids, this research is concerned with the important variables in determining the process and outcome of its use in problem-solving. Under the proposition that problem-solving, with cognitive fit, results in increased problem-solving efficiency and effectiveness, a theoretical framework is needed to provide answers to the research questions: (1) do hypertext-based learning aids enhance the acquisition of advanced knowledge? (2) will the measures of performance (effectiveness and efficiency) increase as a result of learning with a hypertext-based learning aid? (3) will these advantages, effectiveness and efficiency, of using hypertext-based learning aids increase when we increase the level of complexity of the material?

The previous literature review has included a variety of theories that can be applied to this research, but none of them provides a unifying comprehensive framework for this research. Since we are looking to make an assessment regarding tasks and we lack a theory of tasks, we need to find other ways to introduce task complexity as a variable in these theories. The cognitive fit paradigm can be extended to include the level of task complexity. Vessey (91) introduced the nature of the task in her research via the paradigm of cognitive fit. Within this general paradigm, the characteristics of tasks were examined to explain performance in specific problem-solving situations.

3.1 The Theory of Cognitive Fit

Cognitive fit is the notion that problem-solving elements — problem representations, methods, and tools — should support the strategies required to perform a task. If we apply this theory to learning, the outcomes of learning can be used to verify whether the representation of the learning material supports the task of learning it. If that is the case, cognitive fit leads to an effective and efficient problem solution. Otherwise, it does not mean that the task cannot be performed; it simply means that the task outcome is less effective, less efficient or both (i.e., less accurate and/or slower). This theory provides a strong approach to problem-solving, because it is based on matching solution approaches to the problem at hand.

Unfortunately in the literature there is no coherent way to articulate the relationship between the type of task and the problem representation that it facilitates. So far, the theory of cognitive fit has dealt with fairly simple information acquisition and information evaluation tasks. Vessey (91) commenting on the future directions of research on cognitive fit indicated that it "*might address the effects on performance of problem-solving in mismatched contexts, the relative importance of the problem representation, and the nature of problem-solving in more complex tasks.*" She also commented on two possible ways the theory can be extended:

> "*[The theory of cognitive fit] can be applied to any domain where there is sufficient information to permit analysis of the tasks to be performed. [...] It can be extended to include problem-solving techniques, tools, or aids*" (Vessey, 91).

The cognitive fit paradigm is a good example that a general theory of tasks is not essential to develop a problem-solving paradigm. Since theories can be developed for particular research contexts, we can tailor the theory of cognitive fit to examine the effects of information presentation on a more complex task, such as learning. This can be accomplished if we base this approach in its reference discipline: Education. Therefore, educational measurement theories will be used to examine performance on more complex tasks, for analysing the advantages of hypertext over linear text, complemented with other cognitive learning theories that are related to the use of hypertext for learning. Such a theoretical perspective reflects the focus of this research which is the use of technology for problem representation and the "match" or "fit" between that technology and the task of learning from it.

3.1.1 Independent Variable

Problem representation, in the graphs versus tables literature, refers to the way the data are presented, either as a graph or as a table, to decision makers as problem-solving aids. The literature on human information processing theory regarding problem isomorphs provides considerable evidence to support the notion that the processes used by problem-solvers when solving a problem are specific to the problem representation.

Simon and Hayes (76) showed that subjects constructed different mental representations for structurally similar problems (isomorphs), i.e., they derived the mental representation that was more readily available from the problem representation. They also reported that those subjects selected problem-solving processes that were compatible with their mental representation. Accordingly, research on problem isomorphs suggests that subjects perform better when their problem-solving processes are adapted to the problem representation.

Other disciplines that suggest a close relationship between problem representation and problem-solving processes are Behavioural Decision-Making, and Consumer Behaviour. Within the former, Bettman and Kakkar (77) found that even though subjects process information in whatever way they chose, they tend to do so in ways consistent with the representation of the information. So it seems that decision-makers adapt their processing strategies to the problem representation.

Finally, there is also evidence, from the consumer behaviour literature, that matching the problem representation to the task has significant effects on decision-making performance. For example, Bettman and Zins (79) found a difference in the time but not in the accuracy of performance, suggesting that subjects adapt the time to complete a task while keeping accuracy constant. They tested the effects of alternatives, attributes, and mixed alternatives and attributes representations on choice tasks facilitated by either alternative or attribute processing approaches. This also indicates that decision-makers perform better when the problem representation matches the task to be performed.

3.1.2 Dependent Variables

The decision outcomes typically investigated in the graphs versus tables research are performance (usually measured as *decision quality*, i.e., accuracy, or another surrogate for accuracy), interpretation accuracy, and confidence in the results obtained, i.e., satisfaction in decision-making. Vessey (91) restricted her analysis to the objective performance variables (time and accuracy) and interpretation accuracy.

To control for time-accuracy tradeoffs that subjects incorporate into their problem-solving strategies, both time and accuracy were assessed. Furthermore, to fully account for the tradeoff, time and accuracy were assessed jointly, disregarding the perception that in this situation it is more important to make the correct decision than to make a quick one.

The other dependent variable, interpretation accuracy, is used frequently in graphs versus tables research. It is a general term encompassing any set of questions used to assess subjects' understanding of the material.

The nature of the task was introduced in Vessey's discussion via the paradigm of cognitive fit. Within this paradigm, the characteristics of the task can be examined to explain performance in specific problem-solving situations. In that way she overcame the lack of a theory of tasks in the literature. Hence, she found a way to introduce task as a variable into her theory. Researchers dealing with tasks view task complexity as an important characteristic but have been unable to make it operational. The literature has not been able to articulate in a coherent way the relationship between the type of task and the problem representation that facilitates it.

The educational outcomes, as a way to examine some complex forms of human performance, have been classified in groups of capabilities. Gagné (85) makes a distinction among the major categories of human performance by looking at (1) intellectual skills, (2) verbal information, (3) cognitive strategies, (4) motor skills, and (5) attitudes. His is a very naturalistic viewpoint, by describing the major kinds of learning outcomes, as types of human capabilities that are learned. His argument is that a serious and comprehensive examination of human learning must take into account all these varieties.

Educators have devoted considerable effort to reduce the ambiguity associated with stating instructional objectives and translating these objectives into relevant test items. They have divided learning outcomes into three non-overlapping domains: cognitive, affective and psychomotor. The first document stemming from this work is known as Bloom's taxonomy of educational objectives (Bloom, 56). It provides six categories for classifying cognitive behaviours: knowledge, comprehension, application, analysis, synthesis, and evaluation.

The cognitive taxonomy has received the most attention from educators because it has been available the longest and because it describes the set of abilities educators are more interested in measuring. A major contribution of the taxonomy is the awareness it has created regarding the intellectual level at which instructional objectives and test items are written. That is, educators who may have written most of their objectives to require simple remembering or recall of information have come to realize that they actually intended for students to understand and apply knowledge. By using the taxonomy to classify objectives, educators can see more readily whether their expectations are appropriate.

3.1.2.1 Bloom's Taxonomy of Educational Objectives

A group of educational researchers developed what is known as Bloom's Taxonomy of Educational Objectives. A taxonomy that (1) provides classification of the goals of our educational system, (2) facilitates the exchange of information about curricula developments and evaluation devices, (3) specifies objectives to facilitate learning experiences and prepare evaluation devices. It consists of explanations of six levels of thinking in the cognitive domain (i.e., Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation). In the forty years of its existence, Bloom's educational learning taxonomy has been widely used in developing instructional and testing material.

Taxonomies are developed, and used, for classification purposes. Also, taxonomies are helpful tools to insure accuracy of communication among different users. They provide a greater degree of precision in the field in which they are grounded. Any of these uses demand a clear understanding of the structure of the taxonomy, its principles of construction, and its organization.

The idea of developing a classification system was suggested at a meeting of college examiners during the 1948 American Psychological Association Convention in Boston. After long discussions, these examiners agreed that a theoretical framework could be achieved through a classification system of the goals of the educational process. This taxonomy was built as an educational, logical, and psychological classification system (i.e., educational distinctions received major consideration, terms are logically defined as precisely as possible and used consistently, complying with psychological principles and theories).

The taxonomy consists of three parts: the cognitive, the affective and the psychomotor domains. The cognitive domain, which is the concern of this research, includes those objectives that deal with the recall or recognition of knowledge and the development of intellectual abilities and skills. It is the domain in which most of the research in curriculum development and evaluation has taken place.

Basically, this taxonomy is concerned with the changes produced in individuals because of educational experiences. It classifies the intended behaviour of students (i.e., the ways in which individuals act, think, or feel as the result of participating in some unit of instruction). It recognizes that the actual behaviour of the students after they have completed the unit of instruction may differ in degree as well as in kind from the intended behaviour specified by the objectives. In other words, the effects of instruction may be such that the students do not learn a given skill to the desired level of perfection; or, for that matter, they may not develop the intended skill at all. This is a matter of grading or evaluating their performance. The emphasis in the taxonomy is to obtain evidence on the extent to which desired or intended behaviours have been learned by the students.

Within the cognitive domain, Bloom's taxonomy includes remembering, reasoning, problem-solving, concept formation, and creative thinking. Obviously the more complex behaviours include the simpler ones —making it difficult to classify them. The main principle of classification comes from the Gestalt point of view where complex behaviours are more than the sum of simpler ones. Then, for consistency in classification, the taxonomy indicates that a particular behaviour should be placed in the most complex class that is appropriate and relevant.

Taxonomies must be validated by demonstrating their consistency with the theoretical views in research findings of the field they attempt to order.

Then, they must be constructed so the order of their terms must correspond to some real order among the phenomena represented by them. This is the basic problem of any taxonomy, to order phenomena in ways that will reveal some of their essential properties and the relationships among them. Bloom's taxonomy solved this problem by assuming that a particular simple behaviour may become integrated with other equally simple behaviours to form a more complex one. Bloom, and the group of researchers that developed the taxonomy, studied a large number of problems occurring in their examinations and found evidence to support this hypothesis (Bloom, 56).

Probably the most common educational objective is the acquisition of knowledge or information. In other words, it is desired that as a result of completing an educational unit, students will be changed with respect to the amount and kind of knowledge they possess. Frequently, knowledge is the primary, sometimes almost the only, educational objective in a curriculum. By knowledge, we mean that students can give evidence that they remember, either by recalling or by recognizing, some idea or phenomenon with which they have had experience in the educational process.

One of the major problems regarding knowledge is figuring out what is knowable, since there are many ways in which something can be said to be known. To a larger extent, knowledge, as the term is used in North American education, depends on some external authority; an expert or a group of experts is the specialist of knowledge. Also, information results from logical tests of consistency either by definition or by some logic of relationship. Finally, some knowledge or information is known as the result of some historical, experiential, or pragmatic test (Lakatos, 70).

Bloom's taxonomy regards knowledge as basic to all other elements it encompasses. Problem-solving and thinking are not carried out in a vacuum. They are based on some kind of knowledge. Knowledge becomes the material of the problem or the yardstick to verify the accuracy and adequacy of the solution.

Other elements in the taxonomy deal with ways knowledge can be used, i.e., how the information can be applied to new situations and problems, how it can be generalized in dealing with new problems and new materials. It is expected that when students encounter a new problem or situation, they will select an appropriate technique for solving it. This has been labelled critical thinking, reflective thinking, and problem-solving; whatever its name, it requires some analysis, or understanding, of the new situation; it requires a background of knowledge, or methods, ready to be used; it also requires some facility to evaluate the appropriate relations between previous experience and the new situation. What follows is a brief description of the six elements of the taxonomy:

3.1.2.1.1 Knowledge

Knowledge involves remembering, either by recall or recognition, of ideas, materials, or phenomena. It also involves remembering methods and processes, or remembering a pattern, structure, or setting.

3.1.2.1.2 Comprehension

Comprehension represents the lowest level of understanding. It refers to a type of understanding or apprehension such that the individual knows what is being expressed and can make use of the material or idea being communicated without necessarily relating it to other material or seeing its fullest implications. Emphasis is on a grasp of meaning, intent or relationship.

3.1.2.1.3 Application

Application deals with the use of abstractions in particular and concrete situations. The abstractions may be general ideas, rules of procedures, or generalized methods. The abstractions may also be technical principles, ideas, or theories that must be remembered and applied. Emphasis is on applying appropriate principles or generalizations.

3.1.2.1.4 Analysis

Analysis involves the breakdown of a communication into its constituent elements or parts such that the relative hierarchy of ideas is made clear and/or the relations between the ideas expressed are made explicit. Such analyses are intended to clarify the communication, to indicate how the communication is organized, and the way in which it manages to convey its effects, as well as its basis and arrangement. Emphasis is on breaking down into constituent parts, detection of relationships among parts, and of the way they are organized.

3.1.2.1.5 Synthesis

Synthesis comprises the putting together of elements and parts to form a whole. This involves the process of working with pieces, parts, elements, etc., and arranging and combining them in a way that constitutes a pattern or structure not clearly there before the student's performance.

3.1.2.1.6 Evaluation

Evaluation consists of judgements about the value of material and methods for given purposes. It includes quantitative and qualitative judgments about the extent to which some material and methods satisfy the criteria, the use of a standard for an appraisal. The criteria may be those determined by the students or those given to them. Emphasis is on values — making qualitative or quantitative judgements with criteria from internal or external sources and standards.

3.2 Other Theories related to the use of Hypertext for Learning

While the previous section discussed the foundations for this study from the perspective of the Cognitive Fit Theory, this section reviews additional theories that deal with the use of Hypermedia-based systems for learning. Among them, the Theory of Hypermedia Interactions (Jonassen and Mandl, 90) includes the *user* or *learner, tasks, goals, interactions*, and *motives* or *purposes* applicable in hypermedia use, the knowledge domain, and the material used, including the intellectual and/or learning activities involved. This model rests heavily on *user characteristics* such as intellectual abilities, learning skills, meta-cognitive abilities, motivation, plans, intentions and objectives, computer literacy in general and hypermedia literacy specifically. This is not congruent with the Theory of Cognitive Fit since the most useful conclusions presented by Vessey (91) are that *"permitting decision makers complete freedom in choosing their own problem representation(s) will not necessarily lead to improved performance,"* and *"problem solvers will not necessarily perform better when the problem representation they use supports their natural problem-solving strategies."*

Hypertext has been used previously in research studies dealing with teaching and learning in different contexts and with different questions about hypermedia and learning. To understand the different approaches that can be taken to understand the impact of hypertext systems on education, we must look at three global development trends:

[1] There is a growing need for collaborative working environments. The current generation of hypertext tools is mainly restricted to fast browsing and flexible linking, leaving a need for new systems to support creative processes.

[2] The trend to use hypermedia systems for study and training has implications not only on the physical storage device to be used but also on the flexibility of students' and trainees' learning.

[3] The process of knowledge dissemination, i.e., training, teaching, learning, will not remain unaffected by these technologies. An important function of new generations of hypertext systems is the support they will give to externalizing the way we think about complex concept domains.

It is interesting to see that the hypertext idea has elicited quite a spectrum of research lines such as: (1) Research dealing with hypertext to assist in complex mental tasks, such as learning and knowledge acquisition. (2) Research dealing with ways of interaction between users and a hypertext system, looking for the optimal. (3) Research regarding user interfaces for navigation and collaboration. (4) Finally, research about the "right" granularity, i.e., detail, for representing knowledge beyond the level of associative links. Briefly we will discuss each one of them.

3.2.1 Hypertext and complex mental tasks

In this particular line of research, we find experiments or field studies on the actual use of hypertext in learning situations, questioning the level of satisfaction, effectiveness, or both. Among the first studies ever published are those by Legget, et al. (90) on using hypertext in order to study information presented at various levels of complexity.

Most of the studies in this line of research have been presented by educational technologists with a major emphasis in learner control. They were looking to optimize the tradeoff between guidance and learner control, assessing different aspects such as disorientation, confusion, and the lack of inherent coaching with the freedom to browse along, and the openness for students to fill in their own lines of thought.

Also, within this line of research there has been a series of studies dealing with hypertext as a learning tool as opposed to an instructional one. Again, educational technologists are interested in assessing whether hypertext is a teaching tool rather than a learning one or vice versa. Specifically, these studies show that learners incorporate new information into existing knowledge structures. In other words, what readers actually learn from textual information is controlled mainly by the activation of their prior knowledge (Jonassen and Grabinger, 90).

This line of research has also dealt with issues concerning the openness versus closedness of educational hypertext systems. Open versus closed hypertext is the extent to which the system can or should be modifiable by the learner (or the learner's teacher). The content in exploratory hypertext systems is usually conceptualized and written by experts in the field under study. Generally, these systems do not allow additions to that content by the student, except in the form of readers' notes.

3.2.2 Hypertext - User Interaction

This line of research isolate functions such as the flexibility of browsing, the effectiveness of navigation tools, or the amount of error made by the user. Here the interest stays mainly at the level of interaction, and not so much at the cognitive level. Researchers here break down the determinants of interaction benchmarks into individual differences among users and different tasks that users may perform when they use hypertext systems.

At this level, the psychology of human-computer interaction (Card, Moran and Newell, 83) guides researchers by advocating high-level concerns and attitudes about user interface design in about ten principles known as the GOMS[4] model. Unfortunately the GOMS model is not applicable to more complex analysis such as problem-solving.

---

[4] This model considers four components: a set of Goals, a set of Operators, a set of Methods, and a set of Selection rules.

Other researchers have stressed the social and psychological elements of the task situation. In fact, they pay attention to the ecology of the user situation, claiming that an organism and its environment are already coupled, as if they resonate and give support or vice versa. With this in mind, they explain the sense of disorientation some users have while working in a hypertext system. The underlying notion of "space" in hypertext, expressing the extended connections and consequences of "being away" from one's starting point, unifies the feeling that one has "moved" and is unable to determine exactly "where" he or she is.

Metaphor-based interfaces acquired their popularity from the belief that they promote effective mental models. Once there is an understanding that a diskette contains "files," and that those files "contain" a stream of information, users can accomplish new tasks by analogy from their understanding. Explicit metaphors to visualize states and transitions by means of network diagrams have become important as programming tools, document browsers, and tools for knowledge elicitation.

3.2.3 Navigation and Collaboration in Hypertext

This line touches the essential principles of hypertext and the match (or lack of it) with the mental process of users. The basic issue here is how to upgrade hypertext to a level of externalized knowledge rather than manage interlinked fragments of information. Navigation is so important since the most predominant problem of consulting large hypertext resources is the risk of disorientation. Navigation in its pure sense deals with the problem of how to reach a specific point or a set of points in an efficient way. It can be assisted by representing the configuration in space-like maps, schemes, and so forth. Additional mechanisms such as backtrack facilities, a history listing, time-stamping, landmarks such as breadcrumb facilities, footprints, are necessary.

The popularity of hypertext lies in its ability to relate separate elements of texts without prescribing a reading sequence. Hypertext offers support for the authoring of complex documents and the processing of ideas, especially in a group or collaborative environment. Regarding collaboration, hypertext can facilitate collaboration between student and teacher, and also among students. Teachers can stimulate students to confront different points of view, while they discuss and explain common aspects in different opinions among students. It is not clear why the term collaboration was preferred to cooperation, since the latter gives a better indication of what educators mean when they want students to work together towards a common goal.

3.2.4 Hypertext Systems Design

Hypertext authors deal with practical problems about length of paragraphs, number and type of links, etc. Researchers have reported that hypertext links focus on the level of explicit rather than implicit structures in text. Implicit structures in texts are much like the network of concepts in the mind of the author, and presumably in the mind of the reader afterwards. They also advise designers that hypertext systems should be more explicit about the semantic implications for readers traversing a link.

Researchers have come to realize the difficulty of defining "optimal" granularity, i.e., the detail, for representing knowledge beyond the level of associative links, is highly dependent on context sensibility in the domain of information. Information that is factual in nature can be taken out of context more easily. Discourse representation, on the other hand, can be destroyed if designers allow readers to "jump" in the middle of an episode. The arguments for choosing an "optimal" granularity in discourse are more problematic. Solutions nowadays are more subjective than theoretical.

CHAPTER 4. RESEARCH MODEL AND HYPOTHESES

4.1 The Research Model

The research model, depicted in Figure 3, is basically the model of cognitive fit extended to include the theoretical framework identified in the previous Chapter and specific research objectives of the present research as discussed in Chapters 1 and the review of the literature in Chapter 2. This study deals with assessing the cognitive fit of a tool, i.e., a hypertext-based learning aid, designed specifically to deliver an Information System Implementation Model. We are looking at the tool to assess whether it will have an impact on the level of understanding of the learning material by its users. We assume that a difference in understanding will make a difference in the way they will apply it in a problem-solving situation. In other words, the issue is whether performance of its users will improve in effectiveness and efficiency in a problem-solving situation, while applying the model learned.

In our model, **Mental Representation** is the way the problem is represented in human working memory, i.e., the way it is understood. This representation is the output of the characteristics of both the tool used to learn the model and the task. According to Vessey (91), Cognitive Fit Theory views problem-solving as an outcome of the relationship between problem representation and problem-solving task. The independent variable **Tool**, captured by means of the treatments, is a special case of the Problem Representation variable used in the cognitive fit paradigm.

Task, refers to the type of assignment used to elicit different levels of performance. These assignments deal with different levels of the taxonomy of educational

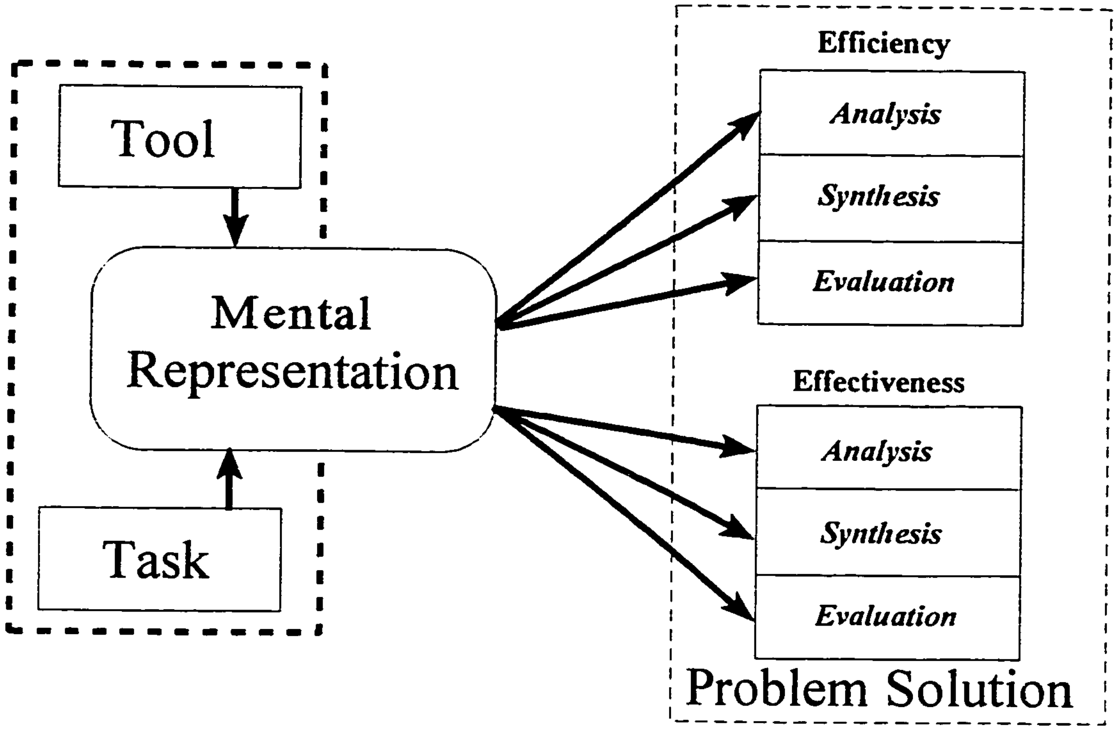


Figure 3 Research Model.

objectives (Bloom, 56). Once learning has happened, the understanding of the material may be elicited from the learners at different levels. It may be elicited as verbal information or some other form, depending on the capabilities targeted: intellectual skills, cognitive strategies, verbal information, attitudes, or motor skills.

Then, the performance that accompanies the learning of a new capability is simply a verification that learning has occurred. Since performance is typically a single act of the learner, it indicates only some reasonable probability that the capability has been acquired in a reliable manner. In order to assure that such stability is indeed present, it is necessary to require additional instances of the performance. Assessing performance in this manner is what is meant usually by "giving a test." The functions served by such a test are twofold, to establish that the newly learned capability has reasonable stability, and to provide additional practice that serves to consolidate what has been learned (Gagné, 85).

With respect to the dependent variable in the research framework of cognitive fit, problem-solving performance has been made operational as **Effectiveness** and **Efficiency** of the **Problem Solution** for each one of the levels of complexity used to design the tasks (Analysis, Synthesis, and Evaluation). **Effectiveness** is captured as a percentage of the maximum score for each level of complexity. So, Effectiveness-in-analysis will represent the percentage of correct answers obtained by a subject working in a multiple choice test. Effectiveness-in-synthesis will represent the percentage of correct inferences about the model learned obtained by a subject working in a short case. Finally, Effectiveness-in-evaluation will represent the percentage of well-supported judgements provided by a subject regarding the information presented in a large case about the implementation of an information system.

Following Vessey's (91) argument that in order to fully account for time and accuracy tradeoffs, both have to be assessed jointly, **Efficiency** is captured as a ratio of the obtained score, divided by the time spent solving the problems for each level of complexity. In the literature, efficiency is measured as speed, but captured by the time spent solving the task, (Vessey, 91, Vessey and Galleta, 91), which we find inappropriate in this learning situation, mostly because we expect to be able to qualify the use of time while learning and solving-problems with the obtained score. So, Efficiency-in-analysis represents the ratio of correct answers obtained by a subject divided by the time spent working in the test. Efficiency-in-synthesis represents the ratio of correct inferences about the model obtained by a subject divided by the time spent working in a short case. Finally, Efficiency-in-evaluation represents the ratio of well-supported judgements provided by a subject in a large case divided by the time spent working on it.

Bloom's taxonomy includes six levels from knowledge to evaluation. We are interested exclusively in the three higher levels of the taxonomy since this research deals with advanced learning. Since the first three levels deal mainly with the knowledge of the "language" of the domain in question, then knowledge, in Bloom's terms involves remembering either by recall or recognition the ideas, material or phenomena of the domain under study. Advanced learning takes these issues as given. Since subjects have had a previous introduction to the theme, they are familiar with these ideas, material or phenomena. Therefore the knowledge of these ideas is independent of the learning method used by subjects in our research.

A similar argument can be made for comprehension, the next level in Bloom's taxonomy, since comprehension represents the lowest level of understanding, with an emphasis on a grasp of meaning or whether learners know what has been expressed and are able to use the material or idea without further explanations. Again, advanced learners are familiar with the terminology of the domain in question to a point that they are able to follow a discussion in the domain without difficulty.

The next level, application, deals with abstractions of elements in the learning domain, mainly to see whether learners are able to identify them in a concrete situation. At this level, subjects are not asked to do more than identify elements in a particular situation. Advanced learners are expected to do that without major problems.

Since analysis involves the breakdown of ideas into its constituent elements, this is the first level of interest in our model. Here, advanced learners need more than an understanding of the material, they are asked to analyse each element and see whether they can indicate how the material is organized, and detect the relationship of its parts. The learning material will definitely have an impact on how these issues are conveyed to the learner. So starting with analysis, and continuing with synthesis and evaluation, these are the levels of complexity of interest in the present research.

4.2 Research Hypotheses

This section presents the hypotheses derived from the research model, in light of the Theory of Cognitive Fit. These hypotheses are organized according to the level of complexity and separated by the variable under study.

Hypotheses regarding effectiveness according to the level of complexity:

H_1: Users learning from the hypertext version of the learning material will perform more effectively on a case study than users learning from the linear version of the same material, when the emphasis of knowledge is on quantitative and qualitative judgments (i.e., level of complexity is Evaluation).

At this level, subjects using the hypertext version are expected to have a deeper understanding of the model and based on that model, make more sound judgments. This means that they will be more accurate in their judgments according to a previously validated instrument.

H_2: Users learning from the hypertext version of the learning material will perform more effectively on a case study than users learning from the linear version of the same material, when the emphasis is on inferences made about the learning material (i.e., level of complexity is Synthesis).

Subjects are expected to make inferences about the model in a specific case situation. Since the understanding of the material will be different from those in the control group from those in the experimental, we expect to observe this difference on the way they make inferences in a case study.

H_3: Users learning from the hypertext version of the learning material will perform as effectively on a multiple choice test as users learning from the linear version of the same material, when the emphasis is on breaking the model into its constituent parts (i.e., level of complexity is Analysis).

At this level, advanced learners will be able to grasp this knowledge from both versions of the material, so they are expected to perform equally well. This hypothesis is presented as a way to validate our approach that the all previous levels of the taxonomy are not really meaningful at the advanced level. In other words, the learning material being implemented in hypertext will not make a difference unless subjects using it are requested to have a deeper understanding of the material, and also to use the information for more complex tasks.

Hypotheses regarding efficiency according to the level of complexity:

H_4: Users learning from the hypertext version of the learning material will perform more efficiently on a case study than users learning from the linear version of the same material, when the emphasis of knowledge is on quantitative and qualitative judgments (i.e., level of complexity is Evaluation).

This hypothesis is similar to H_1. The only difference is the variable of interest. H_4 deals with efficiency in the same way that H_1 dealt with effectiveness. The main difference is that subjects are not only expected to be more accurate, they are also expected to do it faster.

H_5: Users learning from the hypertext version of the learning material will perform more efficiently on a case study than users learning from the linear version of the same material, when the emphasis is on inferences made about the learning material (i.e., level of complexity is Synthesis).

We can present a similar argument between H_2 and H_5 regarding efficiency versus effectiveness. At this level subjects are also expected to perform faster and more accurately.

H_6: Users learning from the hypertext version of the learning material will perform as efficiently on a multiple choice test as users learning from the linear version of the same material, when the emphasis is on breaking the model into its constituent parts (i.e., level of complexity is Analysis).

Finally, the same argument is made between H_3 and H_6 where subjects in both modules are expected to perform equally in effectiveness and in about the same amount of time.

Basically what these hypotheses say is that the effectiveness and efficiency characterising performance of subjects in the experimental group working on problem-solving increases as the level of complexity does, compared to the control group.

These hypotheses are based directly on the Cognitive Fit Theory discussed in Chapter 3. What it is new here, is that by incorporating the level of complexity in this model, we are looking to settle the differences in the results of the use of hypertext for learning literature discussed in Chapter 2, regarding the benefits of using hypertext as a tool for learning.

These hypotheses argue for a reduction of the ambiguity associated with stating instructional objectives and a translation of these objectives into relevant test items. In particular, this research deals with the three highest levels of the cognitive domain: analysis, synthesis and evaluation (Bloom, 56). After forty years of usage, this taxonomy provides a reliable way for designing test material.

CHAPTER 5. RESEARCH METHODOLOGY

This chapter discusses the research methodology including a description of the task domain, experimental material, and experimental design. Also, it presents how the independent and dependent variables were implemented, how subjects were selected and invited, an explanation of the experimental procedures, and how the research instruments were validated.

A laboratory experiment was used to investigate the research questions. This method is taken to be appropriate for the following reasons, among others: The method allows testing the impact of alternative versions of the same learning material while controlling various potentially confounding factors. Experimental systems can be implemented in such a way that they differ only in the way they access the information, while all other aspects are identical including the training and testing material. Addressing a series of concerns generally referred to as tightness of control.

An actual complex IS implementation model was used. Since this model is not regularly taught within the Systems Analysis course, the learning material was developed exclusively for this experiment. This allowed us to address another series of concerns generally referred to as richness of worldly realism.

5.1 Task Domain

A complex and well-supported implementation model was used, the one developed by Lucas, Ginzberg and Schultz (90). It is a rigorous and integrative model of the process of implementation. This model was tested with a generalized Decision Support System in a field study. It combines the two major streams on the research of implementation: process and factor research. The model consists of two stages: the adoption of the system by the manager, and its usage by his or her subordinates. Since each of these two stages is complex in itself, we decided to work exclusively with the manager sub-model (referred to here as the Manager Model).

Lucas, Ginzberg and Schultz (90) developed this structural model of implementation incorporating many of the results of past implementation research studies. It is based on the research traditions of causal modelling, attitude modelling, and innovation process modelling. They used this approach to explain the phenomena of systems implementation with a conceptual model rich in theoretical implications but complex. Its complexity has often discouraged its inclusion in the Information Systems (IS) Curriculum.

The variables included in the model are based on the research findings of previous implementation studies. The model consists of two separable sub-models, the user model and the manager model. The logic of the model is that implementation begins with management initiation and acceptance of a given IS and ends with user satisfaction with the system. Factors[5] leading to manager acceptance are *personal* (MANAGER DECISION STYLE, MANAGER DEMOGRAPHICS), *task-related* (MANAGER JOB CHARACTERISTICS), and *system-specific* (MANAGER KNOWLEDGE OF SYSTEM, MANAGER ASSESSMENT OF SYSTEM AND SUPPORT). TOP MANAGEMENT SUPPORT is seen to influence both MANAGER'S BELIEF IN THE SYSTEMS CONCEPT and MANAGER'S INVOLVEMENT with systems development. Beliefs and

---

5 Factors are listed in SMALL CAPITALS.

involvement do not directly lead to ACCEPTANCE; rather stronger belief leads to more involvement. Also, both stronger belief and more involvement lead to more knowledge of the system. All this to say, that some factors depend on other factors which increases the complexity of the model.

Lucas, Ginzberg and Schultz offered their model primarily to researchers. They see this model as the basis for a "new generation" of research on implementation. We saw the possibility of using it, as our task domain, and at the same time help in its diffusion to the new generation of Information Systems practitioners.

This model was implemented both as computer-based Hypertext and as computer-based Linear text modules. These modules were the experimental material used to deliver the information during the learning session. Both modules were accessed through the same icon on MS Windows 95. What determined which learning module was used, was a log-in code used by the subjects. Without a log-in code, subjects were not able to go beyond the training material. The printouts of both modules are presented in Appendixes A and B.

5.2 Learning an IS Implementation Model

The ultimate objective of developing implementation models is to provide guidelines for the management of implementation, a complex and ill-structured knowledge domain. Research on implementation proceeds from an exploratory stage that sets a conceptual foundation, through the definition of variables and relationships sometimes called dimensions and factors to the integration of these dimensions, factors and their relationships into a testable model.

Most of these models, known mainly by researchers and practitioners, are complex. For some IS professionals, these models are unknown, although these professionals most likely analyse, design and implement information systems. They end up thinking about implementation when they realize, after analysing and designing a system, that the system has to be implemented.

Traditional learning typically treats these categories, or constructs, as regular and well-structured. These models are depicted as taxonomically well-defined (in terms of types, variables and structures). However, the occurrence of these constructs in the real world tends to be more ill-defined, since they may exhibit a variety of characteristics in different situations or in different systems. Definitions often convey some kind of a structure, but in real world practice, events and objects are more ill-structured. Implementation is an ill-structured domain mostly because no general rules or principles exist that can describe or predict most of the cases, there are no defining characteristics to determine appropriate action, and some dimensions are inconsistent between cases, certain aspects of cases are differentially important in different contexts, and each case appears novel because of the interaction of the different factors (Spiro et al., 87, 88).

How can we help instructors concerned about serious misunderstandings of these complex models, if they decide to include them within the learning material in advanced IS courses? Their concerns are justified; they relate their experiences about students' failure to appropriately transfer previously acquired knowledge into more complex issues. Among others, there are two ways to overcome the complexity issue in teaching these models. One is to simplify them by making generalizations or by reducing the number of factors, dimensions or simplifying their relationships. Another is to use a new approach in teaching them by using innovative teaching methods. The present research deals with the second option: the use of Hypertext/Hypermedia as a learning aid to teach these implementation models.

5.3 Research Design

5.3.1 Independent Variable

What follows is a discussion of the independent variable, namely the tool, i.e., the treatments: hypertext and linear text. This variable deals with "representation," one of the variables of the cognitive fit model, or learning method. (Since both versions are computer-based, the use of the computer in itself is not of interest in this setting).

5.3.1.1 The Design of a Cognitive Flexible Hypertext Learning Aid

The manager model was implemented in Hypertext (Figures 4 and 5) by the author under the light of Cognitive Flexibility Theory. This Hypertext module consists of nine screens and ten windows. Each dimension (MANAGER BELIEF IN SYSTEM CONCEPT, MANAGER KNOWLEDGE OF THE SYSTEM, MANAGER ASSESSMENT OF SYSTEM AND SUPPORT, MANAGER INVOLVEMENT, and MANAGER ACCEPTANCE) is presented on a screen. These screens contain their definition and links to the factors or dimensions that have a direct influence on them. Through the links, the windows containing the information of the factors and influencing dimensions are presented when a link is selected. Links are filtered through programs that highlight the relevant information concerning the selected factor or influencing

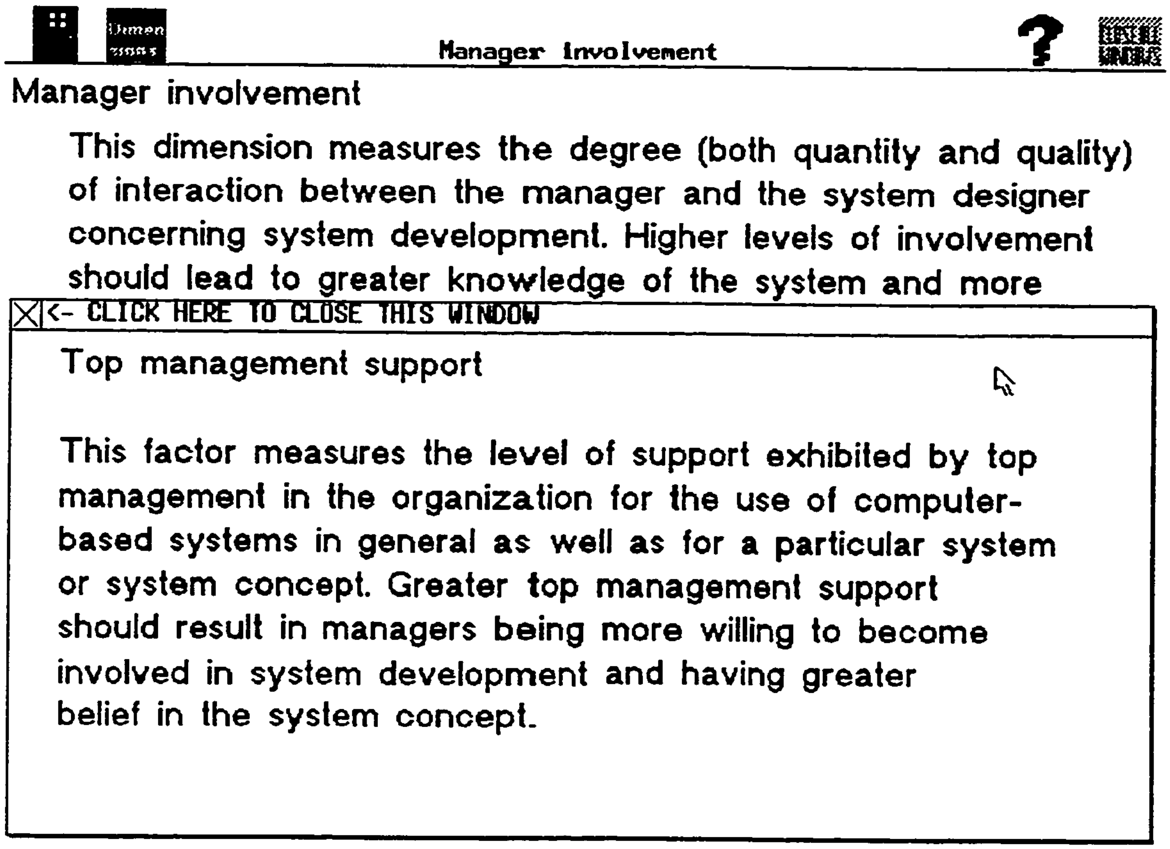


Figure 4 Top Management Support visited from Manager Involvement.

dimension; this is one of the powerful characteristics of hypertext that provides reinforcement in the learning process. The screens also contain linear links through labelled arrows that allow the user to traverse the information sequentially.

This design, besides providing reinforcement in the learning process, directly uses hypertext's cognitive power. As mentioned previously, by showing available links and nodes on each screen, learners visually construct the hyper web which is a representation of the expert's mental model. Also, by revisiting a given factor or dimension, at different times, from different variables with different highlighted information, associations will become apparent. It is a kind of cognitive "click" which allows learners to "know" that

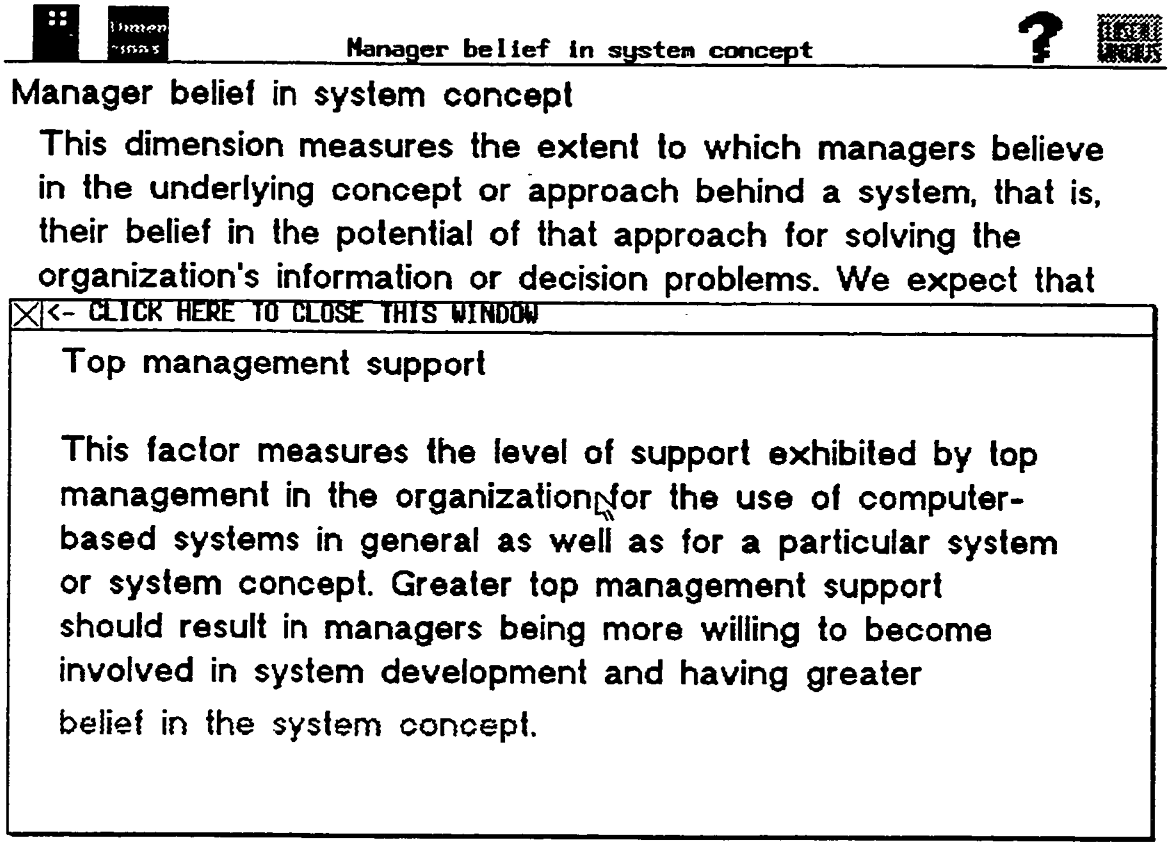


Figure 5 Top Management Support visited from Manager Belief in System Concept.

they have learned something new.

The module also includes four linked icons: DIMENSIONS, CLOSE ALL WINDOWS, HELP, and EXIT DOOR. With the DIMENSIONS icon, users can revisit the dimensions' screen containing a brief explanation of the model and links to each dimension. The CLOSE ALL WINDOWS icon allows users to close all open windows and keep on working in the active screen. HELP opens a window containing a list of the possible actions to take given a particular situation. Students may quit the application using the EXIT DOOR icon.

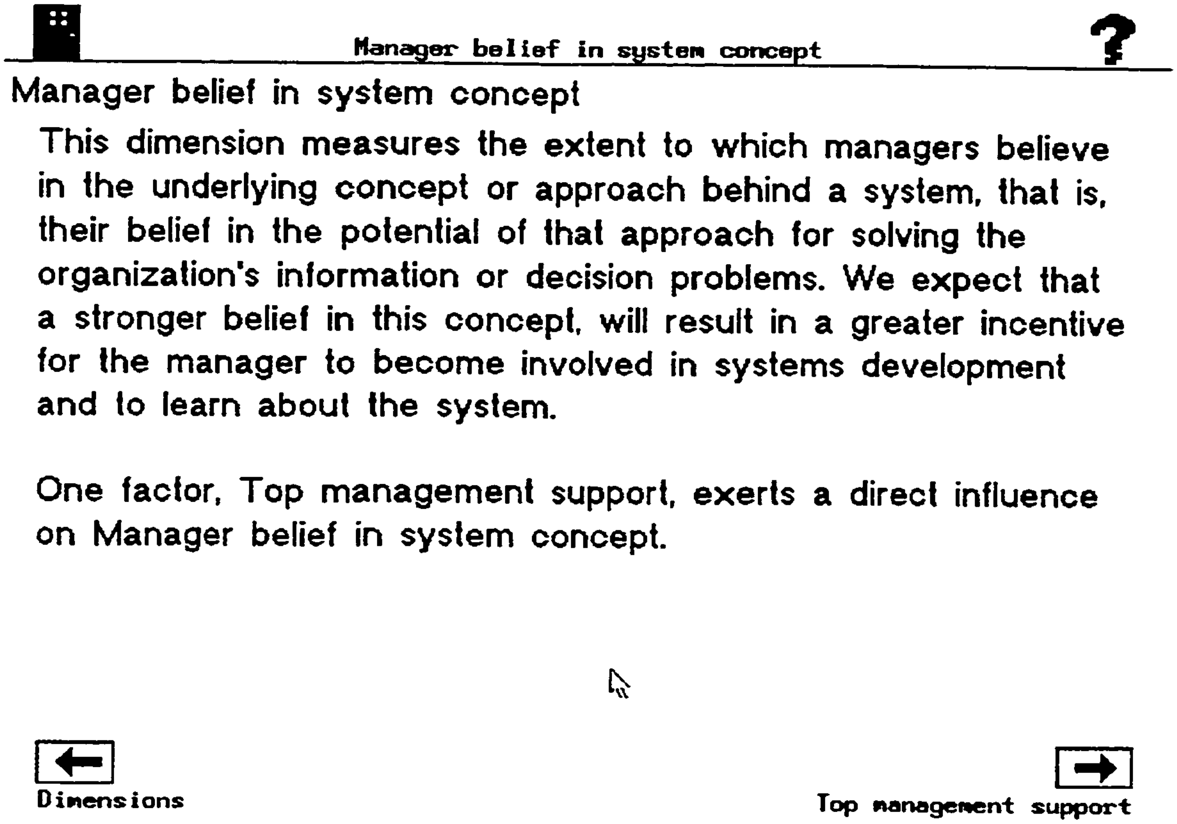


Figure 6: Manager Belief Screen in the Linear Text module.

The main limitation of using hypertext in learning has been determined to be disorientation ("lost in hyper-space.") To prevent this problem, the module contains the DIMENSIONS icon. It takes students to the first screen of the module. Therefore, every time a student needs to revisit the list of dimensions, wants to visit a specific screen, or does not know what to do, by clicking on the DIMENSIONS icon, he or she will be able to get back to a "safe" location. The title of each screen refers to the dimension under discussion. The HELP icon provides information on the options available any time it is requested. With all these aids, disorientation is unlikely.

All these elements are put together to allow learners to experience the model actively instead of passively. Learners can move back and forth within the model to "discover" the relations between dimensions and factors. These relations also help the understanding of the dimensions and factors themselves.

5.3.1.2 The Design of the Linear Text Learning Aid

The linear model consists of fourteen screens. Each factor and dimension is presented independently on a screen. These screens do not contain any links to other factors or dimensions. There are only traversal links through labelled arrows that allow the user to traverse the information sequentially (see Figure 6). The module includes two icons: HELP, and EXIT DOOR. Their functions are exactly the same as their counterpart in the Hypertext version. (See the previous section).

5.3.2 Dependent Variables

The variables of interest were traced by the computer while subjects worked on the learning material (hypertext or linear text), the multiple choice test, and the two case studies. Subjects logged in with a unique code in order to access the learning material. This code identified each one of them during the working session. A tracing program captured the "path" followed, the time a screen or window was opened or closed, the subject's code, and the name of the screen or window. With this information the "time-in-module" was identified and calculated for each user. The tracer also captured the options chosen by the subjects to the questions in the multiple choice test and in the two cases. It also captured the time spent in each one of these three instruments.

The measures of performance, our dependent variable, are effectiveness and efficiency. Since the theory of cognitive fit claims that if cognitive fit exists, problem-solvers' performance will be more effective and efficient, then it can be said that both accuracy and speed will improve. Effectiveness directly measures accuracy, but to control for simple speed-accuracy trade-off, time and accuracy were used in this research and they were evaluated simultaneously as efficiency.

5.3.2.1 Effectiveness in Performance by Level of Complexity

Effectiveness (accuracy) was measured as a percentage of the total score for each level of complexity. Then, Effectiveness-in-Analysis was measured as a percentage of the score obtained by each subject in the multiple choice test. Effectiveness-in-Synthesis was measured as a percentage of correct inferences, based on the learned model, made by each subject working in the first case study. Finally, Effectiveness-in-Evaluation was measured as a percentage of valid judgements, based on the learned model, made by each subject working on the second case study.

5.3.2.2 Efficiency in Performance by Level of Complexity

Efficiency, as mentioned previously, takes into consideration the trade off of speed and accuracy by qualifying it using the following equation:

$$Efficiency = \frac{Score}{Time}$$

$$Efficiency \in [0, \frac{\max(Score)}{\min(Time)}] \qquad \text{Equation } \mathbf{(1)}$$

Since Efficiency generally is expressed as a percentage of an "ideal" performance, i.e., 100% efficiency is never achieved, this variable was transformed to indicate a comparison of the sample. The range of this variable is known, then Efficiency is transformed into New-Efficiency (NEff) using the following transformation

$$NEff = \frac{Efficiency * \min(Time)}{\max(Score)} \qquad \text{Equation } \mathbf{(2)}$$

$$NEff \in [0,\ 1]$$

This new variable was used to measure the comparative efficiency of subjects in each level of complexity. So, Efficiency-in-Analysis ($Neff_{Analysis}$), Efficiency-in-Synthesis ($NEff_{Synthesis}$) and Efficiency-in-Evaluation ($NEff_{Evaluation}$) by each subject were calculated as follows:

$$NEff_{Analysis} = \frac{Score_{Test} * Min(Time_{Test})}{Time_{Test} * Max(Score_{Test})} \qquad \text{Equation } \mathbf{(3)}$$

$$NEff_{Synthesis} = \frac{Score_{CaseI} * Min(Time_{CaseI})}{Time_{CaseI} * Max(Score_{CaseI})} \qquad \text{Equation } \mathbf{(4)}$$

$$NEff_{Evaluation} = \frac{Score_{Case2} * Min(Time_{Case2})}{Time_{Case2} * Max(Score_{Case2})} \quad \text{Equation (5)}$$

5.3.2.3 Learning Time

The third dependent variable, is the "time-in-the-module" or Learning time. This variable is known to have an association with our dependent variables. So it is used as a co-variate to better explain the relationship between the independent and dependent variables (by minimizing error variance). It is measured directly from the tracing program, as the difference between the time when the last window or screen in the learning module was closed and the time when the subject logged in the module.

5.3.3 Experimental Design

A completely randomized design was used. In this study, all participants were randomly assigned to one of the two learning aids, Hypertext and Linear text. To reduce the variance of the error terms in the study, i.e., to make the analysis more precise, we included a concomitant variable or a covariate: Learning Time. This arrangement gives us three possible comparisons for effectiveness and three for efficiency.

With outcomes for effectiveness and efficiency coded $\mathbf{HE_1}$, Hypertext version-Effectiveness in evaluation, $\mathbf{HS_1}$, Hypertext version-Effectiveness in synthesis, $\mathbf{HA_1}$, Hypertext version-Effectiveness in analysis, $\mathbf{LE_1}$, Linear version-Effectiveness in evaluation, $\mathbf{LS_1}$, Linear version-Effectiveness in synthesis, $\mathbf{LA_1}$, Linear version-Effectiveness in analysis, $\mathbf{HE_2}$, Hypertext version-Efficiency in evaluation, $\mathbf{HS_2}$, Hypertext version-Efficiency in synthesis, $\mathbf{HA_2}$, Hypertext version-Efficiency in analysis, $\mathbf{LE_2}$, Linear version-Efficiency in evaluation, $\mathbf{LS_2}$, Linear version-Efficiency in synthesis, and $\mathbf{LA_2}$, Linear version-Efficiency in analysis. Comparisons regarding effectiveness and efficiency outcomes are calculated and tested (see Table 3).

Effectiveness			*Efficiency*	
Hypertext	Linear		Hypertext	Linear
$\mathbf{HE_1}$	$\mathbf{LE_1}$		$\mathbf{HE_2}$	$\mathbf{LE_2}$
$\mathbf{HS_1}$	$\mathbf{LS_1}$		$\mathbf{HS_2}$	$\mathbf{LS_2}$
$\mathbf{HA_1}$	$\mathbf{LA_1}$		$\mathbf{HA_2}$	$\mathbf{LA_2}$

Table 3: Experimental Design.

With this design, all the research hypotheses were tested in a between-subjects manner. Subjects working on the Linear text version were used as a control group for the investigation of the impact of the hypertext-based learning aid.

5.3.3.1 Sample Size

The sample size for the experiment was obtained using power analysis (Cohen, 88). The minimum number of subjects in each cell can be calculated by assuming the research will detect a large main effect, the difference in performance between the two groups will be at least 40% ($f=0.40$), for a significant level of 95% ($\alpha=0.05$) and a power level of 80% ($\beta=0.20$). Since all the subjects worked on the three levels of complexity, the calculation shows that for this factorial design and the parameters previously specified, the number of subjects in each cell (per method) must be at least twenty-one, that means that the minimum number of subjects should be forty-two.

5.4 Subjects

Subjects were one hundred and three undergraduate students (49.51% females, 50.49% males) majoring in MIS in their last year of a Bachelor of Commerce program, enrolled in an Information Systems Analysis course. At this level, the learning material is more relevant and motivated participation. Participation was voluntary. A couple of incentives were offered. Five extra points, participation points, in the course mark were offered to increase participation, and cash prizes to top performers in each group were offered to motivate performance.

5.5 Experimental Procedures

5.5.1 The pre experimental phase

There was an information session for all the students in the targeted population. In this session, subjects were invited to participate in and informed about the study. Participants were informed about the two versions of the learning material, also that they would be randomly assigned to one of those groups, and could spend as much time as they like working with the material. They knew that they were expected to spend some time, between sixty and ninety minutes.

At the end of the session, subjects filled out a consent form indicating their name, ID number, e-mail address and two preferred time-slots when it was more convenient for them to participate in the study. Then, they received a handout, as a reminder, indicating the research schedule and location. (The sign up form and the handout sheet are presented in Appendix F).

A working schedule was prepared. Subjects were informed via E-mail of the time and location of their appointment. One working day previous to their appointment, subjects received a reminder, in the form of an e-mail message. This worked very well. We had 100 percent of subjects participating, and these subjects expressed their satisfaction in the way information was handled using e-mails. The working schedule was set on weekdays, from September 29 to October 10, 1997. Working from 10:00 to 22:00 hours, generated seven daily time slots: Entry times 10:00, 11:30, 14:00, 15:30, 17:30, 19:00 and 20:30 hours respectively.

5.5.2 *Experimental Phase*

The research site was located within the Faculty of Commerce and Administration of Concordia University, for its convenience, since all subjects are students within that faculty. The site contained four desktop microcomputers, PC types, with a 66-mhz Intel® i486 microprocessor, 8 Mega Bytes of Random Access Memory, Super VGA, colour monitors, PS-like mice, running Microsoft® Windows™ 95. The research module was installed in each one of them. A shortcut was created to access the module. This shortcut was labelled RESEARCH and was displayed on Windows™ 95's desktop.

With four microcomputers, seven daily time slots, and ten working days, there were two hundred and eighty working-slots available to accommodate all the subjects. The assignment of subjects was made randomly. One hundred and twenty log-in-codes were generated, sixty for the hypertext module and sixty for the linear text.

Subjects had a hands-on training session. Even though it was almost certain that subjects have had a previous exposure to similar tools: mice, icons, windows, links, etc., their knowledge was not taken for granted. Subjects had the opportunity to work on the training module by clicking on the icon RESEARCH. All the modules were generated on Authority™, a hypertext authoring system developed by *Interactive Image Technologies, Ltd.*, a Canadian software company based in Toronto. Authority™ is a C-based (visual programming) tool for building interactive hypermedia applications. Most of the programming in Authority™ is made with AuthorLang®, a VisualBasic-type of language. These modules were developed by the author of this thesis.

The first module, training, consists of six screens and nine windows. A printout of the training module is presented in Appendix C. The purpose of this module is to allow subjects to move successfully around four locations (HomePage, Point A, Point B, and Point C) while reviewing information. Subjects are welcomed into the training module with an open window containing the instructions. After reading that information, users close that window and visit the first screen, the title screen. When they click on the only available link (a labelled button) they visit a screen labelled "Home Page!" In this screen they are introduced to icons and links, for those subjects that have not had previous exposure to these concepts. They are invited to try an icon, if they do so, they visit a screen labelled "Point C" which is only a way to show them how icons work. The only option there is to click on the icon and return to "Home Page!" After that they can visit either "Point A" or "Point B." "Point A" is a screen that explains and shows how to close open windows, one at a time, or all open windows simultaneously by using the icon labelled "CLOSE ALL WINDOWS." "Point

B" is a screen that explains how to open windows by clicking on highlighted links and see how the links are colour updated after using them (these features are commonly used by Web-users). After they have tried those links, they may exit the training module, or review it again. To do so, they can follow highlighted links or use another icon labelled "EXIT."

5.5.2.1 Experimental Task

When subjects were familiar with the system's features, they were exposed to the learning material. Students were allowed to spend as much time as needed learning its content. The learning module, contains the information of Lucas, Ginzberg and Schultz' (90) Implementation Model. There were two versions of this material, a hypertext-based, and a linear version. To access them, subjects were welcomed into a title screen. By clicking on the only available link (a button labelled: THE MANAGER MODEL), subjects were prompted for their log-in-code (a four-digit password). Subjects received this password when they arrived at the research location. That code was the key to access the hypertext or the linear text version of the Implementation Model.

5.5.2.2 The Hypertext Module

As mentioned previously in section 5.3.1.1, the Hypertext module consists of nine screens and ten windows. The first screen is an introduction to the model and the learning objectives are stated as goals.

Subjects go through the remaining screens and windows at their leisure. Every time they move from one screen to another (or to a window), a tracer captures the time they access that location and writes this information into a tracing file. All this is transparent to the user. Links are filtered through programs that highlight the relevant information concerning the selected factor or influencing dimension. The screens also contain linear links through labelled arrows that allow the user to traverse the information sequentially. It really does not matter the way they traverse all the information, the tracer captures this "personalized" trip and keeps the information for future analysis.

5.5.2.3 The Linear text Module

Subjects using the Linear text module, as described previously in section 5.3.1.2, are also prompted with an introduction to the model and the learning objectives, stated as goals. Subjects go through the remaining screens and windows at their leisure through sequential links. Every time they move from one screen to the next, they have the option of going backward or forward. The tracer captures this information as well as the time they accessed that location and writes this information into a tracing file. As in the hypertext module, all this is transparent to the user.

Subjects received a notepad and were told to use it, if they needed, to take notes while they were working on the learning material. Students were allowed to use their notes while working on the multiple choice test and the two case studies. At the end of the session, these notes were handed in and kept as part of the working material along with the sign up forms for each individual.

When subjects finished working with their version (hypertext, linear text) of the learning material, they were advised that they had finished with it. They had two options in the form of labelled links (buttons: REVIEW MODULE, and WRITE THE TEST). When they chose to write the test, they left the learning module and moved on to the test and the cases.

5.5.2.4 The Test and the Cases

The testing material consists of three parts: A multiple choice test dealing with analysis of the information presented on the learning module; A short case-study, in which the subjects are expected to make some inferences based on the learning material; A longer case-study dealing with the implementation of an information system, subjects are expected to make judgments regarding the success (failure) of its implementation, based on the learning material. These testing materials were validated through a pilot study. (All these materials are presented in Appendix D).

5.5.3 The post experimental phase

After they had finished with the second case study, subjects were asked to give their opinion regarding their experience. The way they gave their opinion, was by using a computer-based evaluation form consisting of twelve Likert scale keys associated in six factors: PARTICIPATION, CONTENT, TASK, TEST, HARDWARE, and GLOBAL SATISFACTION. Three of these factors had some sub-factors. Factor CONTENT included MOTIVATION, USEFULNESS, and EXPERIENCE. Factor TASK includes MOTIVATION, SATISFACTION, and USER FRIENDLINESS. Factor TEST includes GRADE OF DIFFICULTY, CHALLENGE, and MOTIVATION. (This form is presented in Appendix E).

The traced information for each participant was saved into a file. These files were named using the participant's code. A report was generated for each file. These reports generated all the data needed for the analysis. Also, a report for each student was generated and sent via e-mail. Those e-mails contained information regarding their performance and listed the names of the top performers in each group. Prizes were granted accordingly.

5.6 Validating the Tool: The Pilot Study

Subjects in the pilot study were thirty-six (58.33% males, 41.67% females) undergraduate students, majoring in Management Information Systems, enrolled in an Information Systems Analysis course. Participation was voluntary and evaluated as a Lab Project. It represented 5% of the course grade. We had 72% participation. Those students that did not participate were assigned another lab project of equal value. Subjects were randomly assigned to one of two experimental groups: L (Linear) and H (Hypertext).

5.6.1 Procedure

Subjects in the pilot study followed exactly the same procedures as those described previously as *Experimental Procedures* in Section 5.5. Data obtained from this pilot study include the time-in-module, or time spent learning the module, the time-in-test, time-in-case-1, and time-in-case-2, the option selected for each question, the scores in analysis, synthesis and evaluation, i.e., each level of complexity, and a total score for the test. These data were kept in a record, including information of the task type (hypertext/linear), for each subject.

The term validity, when applied to educational material, refers to the precision with which the material measures some cognitive ability. There are two aspects of validity: what is measured and how precisely it is measured. The cognitive abilities referred to are abilities to perform observable tasks. How precisely a test measures an ability is indicated by the reliability of the scores. Reliability is a necessary condition of validity, though it is not a sufficient condition (Ebel and Frisbie, 86). Reliability of the test was assessed with an item difficulty and discrimination analysis.

5.6.2 Pilot Study's Results

The data obtained in the pilot study is presented in Appendix G. It includes the case summaries in Learning Time (Time-in-Module), Effectiveness-in-Analysis, Effectiveness-in-Synthesis, Effectiveness-in-Evaluation, Efficiency-in-Analysis, Efficiency-in-Synthesis, and Efficiency-in-Evaluation, and data obtained with the twelve Likert keys of the evaluation, for each of the two groups: Hypertext and Linear text, as well as for the complete sample.

The elicitation of the performance that reflects the newly learned model seems a reasonable natural event. Performance that exhibits what has been learned provides evidence to the extent to which learning has attained its objectives. Many techniques of assessment by means of tests have been discussed in the educational literature, with their pros and cons.

Even though there exists a great controversy over testing, so far we have no other paradigm for assessing the outcomes of learning as accepted as Educational

Tests. When people learn, we assume there is some internal store of skill and information they are building and adding to it. This internal change results in changes in abilities. When a test is given, we want answers to the test questions to reflect how much of the ability the person possesses. Since we cannot see the ability directly, the ability is sometimes referred to as a latent trait. The goal of Measurement Theory is to describe by means of mathematical functions, the relationship between what can be observed (the item response) and the unobservable (the trait).

These mathematical models (functions) are based on assumptions about the test data; each model makes somewhat different assumptions about them. When a teacher tells a student that she is at the ninetieth percentile in math ability, she also needs to know the group she is being compared to, and the test used has to be specified to make sense of the teacher's statement. Without this additional information, such an assessment is only a subjective evaluation.

A quite different approach is possible, one in which no assumptions need to be made about the ability distribution of the people used. This new approach assumes instead a very simple model that says that the outcome in a test is solely controlled by the difference between the ability of the individual and the difficulty of the items; Nothing more. The more able the person, the better her chances for success with any item. The easier the item, the more likely any person is to solve it. It could be as simple as that. This is precisely the Rasch Model (Rasch, 60, 61; Wright and Stone, 79).

5.6.2.1 The Rasch Model

The Rasch Model is a mathematical model that combines the parameter β_v, which represents person ν's ability β, and δ_ι, for item ι's difficulty δ. Their difference $(\beta_v - \delta_\iota)$ governs the probability of what is supposed to happen when person ν uses his ability β against the difficulty δ of item ι. Then we can measure the probability of a success (x = 1) or a failure (x = 0) using the following equation.

$$P\{ x_{v\iota} = x | \beta_v, \delta_\iota \} = \frac{e^{x(\beta_v - \delta_\iota)}}{1 + e^{(\beta_v - \delta_\iota)}} \quad \text{Equation (6)}$$

Since we do not know the real values of β_v or δ_ι, we can estimate these values by making some assumptions. The Rasch model assumes that the persons' abilities (β_v) are more or less normally distributed with mean M and standard deviation σ and that the items' difficulties (δ_ι) are also more or less normally distributed with average difficulty H and difficulty standard deviation ω.

If $\beta_v \sim N(M, \sigma^2)$ and $\delta\iota \sim N(H, \omega^2)$, then in a sample of size N, working on a test of L items, we can observe the individual scores r_v, and the item scores s_ι. These s_ι values represent the number of people that correctly responded to each item. It follows that the likelihood estimators are b_v and d_ι, defined as: [In the following set of equations, U represents the standard deviation of the items' difficulties, and V represents the standard deviation of the individuals' scores. The values $2.89 = 1.7^2$, and $8.35 = 1.7^4$, come from the scaling factor 1.7 which brings this ogive into approximate coincidence with the normal ogive (Rasch, 60)].

$$b_v = H + \sqrt{\left(1+\frac{U}{2.89}\right)\left(1-\frac{UV}{8.35}\right)}\; Ln\left[\frac{r_v}{(L-r_v)}\right]$$

$$d_\iota = M + \sqrt{\left(1+\frac{V}{2.89}\right)\left(1-\frac{UV}{8.35}\right)}\; Ln\left[\frac{(N-s_\iota)}{s_\iota}\right]$$

Equation **(7)**

The estimate of the probability of person v answering item ι (either correctly x = 1, or incorrectly x = 0) is $p_{v\iota}$ calculated as

$$P_{v\iota} = \frac{e^{x(b_v - d_\iota)}}{1 + e^{(b_v - d_\iota)}}$$

Equation **(8)**

We can use these probabilities to calculate the sum of their squared residuals $z_{v\iota}^2$ for each item, and evaluate item fit. According to the Rasch model this sum of squared normal deviates ($C_\iota^2 = \Sigma\, z_{v\iota}^2$) approximates a chi-squared distribution with $df_\iota = (N-1)(L-1)/L$ degrees of freedom. We can use $v_\iota = C_\iota^2 / df_\iota \sim F_{df\iota,\infty}$ to test the hypothesis that an item fits or not in the test.

In our pilot study, a set of thirteen items (L = 13) was administered to a sample of thirty-six subjects (N = 36). We had a 36X13 matrix of responses $x_{i,j}$ for which i = 1 to 36, and j = 1 to 13. Using the marginal sums of rows and columns as the subjects scores $r_i = \Sigma\, x_{i,j}$, and the item scores $s_j = \Sigma\, x_{i,j}$, we calculated the item difficulties d_j. These difficulties, presented in Table 4, indicate the proportion of incorrect responses for a particular item. Once the parameters were estimated, we calculated the matrix of probabilities $p_{i,j}$ using equation (8). These probabilities were used to calculate the standardized residuals $Z_{i,j}$ using equation (9)

$$Z_{i,j} = \frac{(x_{i,j} - p_{i,j})}{\sqrt{p_{i,j}(1 - p_{i,j})}} \qquad \text{Equation (9)}$$

The sum of squared residuals $Z_{i,j}^2$ as mentioned previously can be used to evaluate item fit. The statistic $v_j = \Sigma Z_{i,j}^2 / df_j$ is distributed as an F distribution with df_j and ∞ degrees of freedom. This F-test is presented in Table 4.

Item	Difficulty	F-test	p-value
Q1	0.194	0.844	0.694
Q2	0.250	0.807	0.742
Q3	0.361	0.935	0.572
Q4	0.417	0.609	0.935
Q5	0.361	0.743	0.818
Q6	0.500	1.459	0.103
Q7	0.694	1.559	0.069
Q8	0.361	1.103	0.364
Q9	0.528	1.212	0.256
Q10	0.139	0.851	0.686
Q11	0.639	1.244	0.229
Q12	0.667	1.821	0.023*
Q13	0.361	1.207	0.260

Reject item at*α=0.05, One-tailed test, df_1=32.3, df_2=∞

Table 4: Difficulty Analysis using the Rasch Model

Discrimination, the other important value needed in the difficulty and discrimination analysis, is known as the point bi-serial correlation between item response and post test score. The point bi-serial correlation gives a measure of association between a true dichotomous variable and a second variable which is measured on an interval or ratio scale. It measures the relationship between performance on each item and on the entire test. In this pilot study, the dichotomous variable was the item response and the second variable was the post test score. For a particular test item, the dichotomous variable was assigned a value of zero (x = 0) for an incorrect response, and one (x = 1) for a correct response. The point bi-serial correlations for each item were calculated in SPSS and t-tests were used to determine their statistical significance. These results are presented in Table 5.

Item	Discrimination	t-test	p-value
Q1	0.258	1.559	0.128
Q2	0.404	2.587	0.014*
Q3	0.384	2.417	0.021*
Q4	0.665	3.836	0.000**
Q5	0.541	3.591	0.001**
Q6	0.183	1.088	0.284
Q7	0.094	0.553	0.584
Q8	0.462	2.996	0.005**
Q9	0.487	3.190	0.003**
Q10	0.479	3.132	0.003**
Q11	0.360	2.248	0.031*
Q12	0.020	0.116	0.908
Q13	0.394	2.506	0.017*

Significant at *α=0.05, **α=0.01, two-tailed test, n = 36, df = 35

Table 5: Post-Test Item Analysis

The discrimination values indicate that the test items discriminated between the good and not so good students. For example, on question four (**Q4**), its discrimination of 0.665 indicates the strongest relationship between student responses and post test scores. In this particular question, those missing the item tended to score low while those responding correctly tended to score high on the test.

If we compare both analyses, we can see that question twelve (**Q12**) is shown as problematic in both cases. The Rasch model rejects this item exclusively while the point bi-serial correlation indicates some problems with items one, six, seven and twelve (**Q1, Q6, Q7, Q12**). In contrast with the point bi-serial correlation, the Rasch item mean square residual has a useful statistical reference distribution. The reference value for testing the statistical hypothesis that an item belongs to a test is a mean square of one with a standard error of $(2/f)^{1/2}$ for f degrees of freedom. Then, the extent to which an observed mean square exceeds the expected value of one can be tested for its statistical significance at whatever significance level is considered useful. Based on this argument, we decided to keep all items but **Q12**. Question twelve was reworded as a result.

CHAPTER 6. QUANTITATIVE DATA ANALYSES

The results of the data analyses are reported in this and the next chapter. Each of these chapters concentrates on one of two distinct types of data, i.e., quantitative measures of process and outcome variables, and qualitative verbal protocol data. This chapter presents statistical analyses on performance measures such as Effectiveness and Efficiency. The next chapter reports on verbal protocol data.

A variety of statistical analysis tools were used to analyse the results obtained in the experiment described in the previous chapter. The selection of the statistical tools was based on the nature of the data, hypotheses tested, and underlying assumptions of the statistical models used. In order to test the hypotheses presented in Chapter 4, an Analysis of Co-variance (ANCOVA) model was constructed after the data were screened to verify that the assumptions imposed by this model were satisfied.

6.1 Data Screening

In Appendix H the descriptive statistics of the individual variables are presented. Table H1, includes information on the overall sample size (N=103) while Tables H2 and H3 contain information about the two groups, Hypertext and Linear text respectively. Descriptive statistics include the mean, standard deviation, minimum and maximum values for all the variables: Learning Time, Time-in-Analysis, Time-in-Synthesis, Time-in-Evaluation, Effectiveness-in-Analysis, Effectiveness-in-Synthesis, Effectiveness-in-Evaluation, Participation, Content (Motivation, Usefulness, Experience), Task (Motivation, Satisfaction, User Friendliness), Test (Difficulty, Challenge, Motivation), Hardware Use, and Global Satisfaction. A summary of these results is presented in Table 6.

Variable		**Method**	**N**	**Mean**	**S. D.**
Learning Time		Hypertext	52	21:37	10:30
		Linear Text	51	19:59	7:01
		Total	103	20:48	8:56
Effectiveness	Analysis	Hypertext	52	6.27	2.84
		Linear Text	51	5.73	2.59
		Total	103	6.00	2.72
	Synthesis	Hypertext	52	2.31	1.94
		Linear Text	51	2.53	1.93
		Total	103	2.42	1.93
	Evaluation	Hypertext	52	11.58	4.41
		Linear Text	51	9.41	4.55
		Total	103	10.50	4.59
Efficiency	Analysis	Hypertext	52	56.93%	25.70%
		Linear Text	51	51.72%	23.22%
		Total	103	54.35%	24.52%
	Synthesis	Hypertext	52	35.54%	30.78%
		Linear Text	51	38.86%	29.94%
		Total	103	37.18%	30.27%
	Evaluation	Hypertext	52	55.03%	19.87%
		Linear Text	51	45.23%	22.09%
		Total	103	50.17%	21.47%

Table 6: Summary of Descriptive Statistics by Method

In Appendix I, a Normal Probability Plot and a Detrended Normal Plot are presented for each of the variables of interest. Since the distributions of the four time variables (Learning time, Time-in-Analysis, Time-in-Synthesis and Time-in-Evaluation) were quite skewed, some transformations were considered.

6.2 Data Transformations

Each one of the time variables was transformed using a Natural Logarithm transformation. These transformations worked well, as it can be seen in the Normal probability Plots and Detrended normal plots presented in Appendix I. Figures I9 and I10 for Learning Time. Figures I11 and I12 for Time in the multiple-choice test, i.e., time in analysis. Figures I13 and I14 for Time in Case 1, i.e., time in synthesis. Figures I15 and I16 for Time in Case 2, i.e., time in evaluation.

After the transformations, One-sample Kolmogorov-Smirnov Tests were conducted for the transformed time variables. These tests are presented in Table 7. The null hypotheses that these variables are normally distributed cannot be rejected ($\alpha = 0.05$). The largest Kolmogorov-Smirnov statistic (K-S Z) is 0.947 for the Natural Logarithm transformation of Time-in-Case-1 (Synthesis) with a p-value of 0.331.

	LN Learning Time	LN T-in-Test	LN T-in-Case-1	LN T-in-Case-2
N	103	103	103	103
Mean	7.0414	6.0260	5.7439	6.8091
Std. Deviation	0.4290	0.3521	0.4535	0.3891
K-S Z	0.539	0.571	0.947	0.583
Sig. (2-tailed)	0.934	0.900	0.331	0.886

Table 7: One Sample Kolmogorov-Smirnov Test

6.3 The Analysis of Covariance

Since there is no reason to use the analysis-of-covariance procedure if the dependent variables are not correlated with the variable considered as co-variate, a correlation matrix was prepared after the transformations to evaluate the association between the variable Learning Time and the dependent variables. This correlation matrix is presented in Appendix J. To test for significant correlations a series of t-tests were run. These tests are presented in Table 8. Learning time is significantly correlated with Effectiveness and Efficiency in Analysis (α=0.01), and with Effectiveness in Evaluation (α=0.05). These tests supported the decision of including the variable Learning Time as a co-variate in our ANOVA model. (Since we are including one co-variate, now we are going to use an ANCOVA model).

Learning Time	Effectiveness			Efficiency		
	Analysis	Synthesis	Evaluation	Analysis	Synthesis	Evaluation
Correlation	0.340	-0.022	0.241	0.288	-0.069	0.170
t-test	3.421**	0.226	2.500*	3.041**	0.698	1.734
p-value	0.000	0.822	0.014	0.003	0.487	0.086

Table 8: Significant Correlations *α=0.05, **α=0.01

In order to test for Homogeneity of Variance, a series of Levene's tests of Equality of Error Variances were performed. Levene's test was chosen because it is less dependent on the assumption of normality than most tests and thus is particularly useful with analysis of variance and covariance. The results of the test are presented in Table 9. Each one of these procedures tests the null hypothesis that the error variance of the dependent variable is equal across groups. The observed p-values indicate that the null hypotheses cannot be rejected. Given the results of these Levene tests, there appears to be no reason to suspect the homogeneity-of-variances assumption.

Variable		Levene Statistic	df1	df2	p-value
Effectiveness	Evaluation	1.101	1	101	0.297
	Synthesis	0.203	1	101	0.653
	Analysis	0.624	1	101	0.431
Efficiency	Evaluation	1.334	1	101	0.251
	Synthesis	0.345	1	101	0.558
	Analysis	0.743	1	101	0.391

Table 9: Levene's Tests of Homogeneity of Variances

Once the preliminary steps of examining the distribution of the variables for outliers, non normality, and inequality of variances have been taken and no significant violations have been found, we can proceed to test our hypotheses.

Even though it seems that the optimal way to test these hypotheses is to test them simultaneously using a MANCOVA model, actually it is not the case in this particular study. A simultaneous analysis will not only test each pair of differences but also other differences that are not of interest in this study, i.e., any linear combination of the dependent variables. This study is particularly interested in a set of pair-wise comparisons that was specified in advance of the data analysis. When that is the case, it is possible to run a series of ANCOVA procedures instead of one MANCOVA.

To analyse our hypotheses, we ran three ANCOVA models for Effectiveness and three for Efficiency, one for each level of complexity. In order to compensate for the deficiencies of running simultaneous tests, the uni-variate analyses were adjusted using the Holm Simultaneous Testing Procedure (Holm, 79).

6.3.1 Testing Hypotheses regarding Effectiveness

Table 10 includes the tests of Between-Subjects Effects for the dependent variable Effectiveness in Evaluation. From this information we can conclude that Hypothesis H_1, is supported (p-value = 0.008). This means that as expected we do find differences in effectiveness between groups working in the two methods (Hypertext, Linear text), when the level of complexity is Evaluation. We can also conclude that the co-variate, Learning Time, helps us explain this difference (p-value = 0.009).

Source	Type III	df	Mean Sq.	F	p-value	Power
Corrected Model	256.327	2	128.164	6.783	0.002	0.912
Intercept	26.842	1	26.842	1.421	0.236	0.219
Learning Time	135.625	1	135.625	7.178	0.009	0.756
Method	113.060	1	113.060	5.984	0.008	0.678
Error	1889.420	100	18.894			
Total	13512.000	103				
Corrected Total	2145.748	102				

Table 10: Test of Between-Subjects Effects of Effectiveness in Evaluation.

Table 11 includes the tests of Between-Subjects Effects for the dependent variable Effectiveness in Synthesis. From this information we can conclude that Hypothesis H_2, is not supported (p-value = 0.281). This means that we do not find differences in effectiveness between groups working in the two methods (Hypertext, Linear text), when the level of complexity is Synthesis. We can also conclude that the co-variate, Learning Time, does not help us at this level of complexity (p-value = 0.907).

Source	Type III	df	Mean Sq.	F	p-value	Power
Corrected Model	1.318	2	0.659	0.174	0.840	0.076
Intercept	1.582	1	1.582	0.419	0.519	0.098
Learning Time	0.051	1	0.051	0.014	0.907	0.052
Method	1.280	1	1.280	0.339	0.281	0.089
Error	377.731	100	3.777			
Total	981.000	103				
Corrected Total	379.049	102				

Table 11: Test of Between-Subjects Effects of Effectiveness in Synthesis.

Table 12 includes the tests of Between-Subjects Effects for the dependent variable Effectiveness in Analysis. From this information we can conclude that Hypothesis H_3, is supported (p-value = 0.335). This means that, as expected, we do not find differences in effectiveness between groups working in the two methods (Hypertext, Linear text), when the level of complexity is Analysis. We can also conclude that the co-variate, Learning Time, helps us explain this lack of difference (p-value = 0.000).

Source	Type III	df	Mean Sq.	F	p-value	Power
Corrected Model	120.491	2	60.245	9.510	0.000	0.977
Intercept	47.922	1	47.922	7.565	0.007	0.778
Learning Time	112.878	1	112.878	17.818	0.000	0.987
Method	5.953	1	5.953	0.940	0.335	0.160
Error	633.509	100	6.335			
Total	4462.000	103				
Corrected Total	754.000	102				

Table 12: Test of Between-Subjects Effects of Effectiveness in Analysis.

6.3.1.1 Parameter Estimates for Effectiveness

Using ANCOVA's mathematical model to express the relationship between the dependent variables, the co-variate, and the independent variables, we proceed to estimate the parameters of the variables in the model. We used the Deviation Contrasts procedure to calculate the parameters presented in Table 13:

Dependent Variable	Parameter	B	Std. Error	t	p-value
Effectiveness Analysis	Intercept	-11.515	4.100	-2.809	0.006*
	Learning Time	2.453	0.581	4.221	0.000*
	Method	0.481	0.496	0.969	0.335
Effectiveness Synthesis	Intercept	2.160	3.166	0.682	0.497
	Learning Time	0.005	0.449	0.117	0.907
	Method	-0.223	0.383	-0.582	0.281
Effectiveness Evaluation	Intercept	-9.486	7.080	-1.340	0.183
	Learning Time	2.689	1.004	2.679	0.009*
	Method	2.096	0.857	2.446	0.008*

Table 13: Parameter Estimates for Efficiency, Method: 1=Hypertext, 0=Linear text.

6.3.1.2 Holm Adjustments for Effectiveness

The Holm simultaneous testing procedure (Holm, 79) is a refinement of the Bonferroni adjustment for conducting a family of tests. It is applicable when the simultaneous tests is a particular set of pair-wise comparisons, contrasts, or linear combinations that is specified by researchers in advance of the data analysis.

We have tested the following three sets of hypotheses:

Test 1: H_o: $\mu_{\text{Effectiveness-in-Evaluation} \mid \text{Hypertext}} \leq \mu_{\text{Effectiveness-in-Evaluation} \mid \text{Linear text}}$

H_a: $\mu_{\text{Effectiveness-in-Evaluation} \mid \text{Hypertext}} > \mu_{\text{Effectiveness-in-Evaluation} \mid \text{Linear text}}$

Test 2: H_o: $\mu_{\text{Effectiveness-in-Synthesis} \mid \text{Hypertext}} \leq \mu_{\text{Effectiveness-in-Synthesis} \mid \text{Linear text}}$

H_a: $\mu_{\text{Effectiveness-in-Synthesis} \mid \text{Hypertext}} > \mu_{\text{Effectiveness-in-Synthesis} \mid \text{Linear text}}$

Test 3: H_o: $\mu_{\text{Effectiveness-in-Analysis} \mid \text{Hypertext}} = \mu_{\text{Effectiveness-in-Analysis} \mid \text{Linear text}}$

H_a: $\mu_{\text{Effectiveness-in-Analysis} \mid \text{Hypertext}} \neq \mu_{\text{Effectiveness-in-Analysis} \mid \text{Linear text}}$

If we apply the standard Bonferroni procedure by obtaining the two-sided P-value of each test statistic and comparing it with $\alpha/3$, we conclude H_0 if the observed P-value equals or exceeds this value ($\alpha/3$), else we conclude H_a.

The Holm procedure carries out the simultaneous testing by obtaining the P-value for each test but then modifies the level against which the P-value is compared in order to improve the power of the test (Holm, 79). Consequently, the Holm procedure may find significant effects when the Bonferroni procedure does not, for the same significance level α. This procedure, slightly more complex computationally, allows adjustments for one side tests, since a P-value must be found for each test.

The first step in this procedure is to rank the set of tests. Rank 1 is assigned to the smallest P-value. The P-value is then compared to $\alpha/3$ to determine whether H_0 or H_a is to be concluded. If H_0 is concluded, the testing is terminated and H_0 is concluded for all other tests. If H_a is concluded for the test with the smallest P-value, the Holm procedure considers next the test with the second-smallest P-value. This P-value is compared to $\alpha/2$. If H_0 is concluded, testing is terminated and H_0 is concluded for the remaining tests. If H_a is concluded, the third P-value is compared to α.

In our experiment we chose a level of significance $\alpha = 0.05$. Then we ranked the three tests according to their p-value. Test 1, with p-value = 0.008 ranks 1, this p-value is compared to $\alpha/3 = 0.05/3 = 0.01667$; H_a is concluded. Then Test 2, with p-value = 0.281 ranks 2. This value is compared to $\alpha/2 = 0.05/2 = 0.025$; H_0 is concluded and the procedure stops. H_0 is concluded for Tests 2 and 3.

Once we have estimated the parameters of MANCOVA's mathematical model for our set of variables and performed the Holm adjustments, we proceed to estimate the 95% confidence intervals for the differences found in Effectiveness in Evaluation. Since the difference in performance, as expected, is found to be higher for subjects in the Hypertext group, the difference is calculated as (Hypertext minus Linear text), which gives us a Mean Difference for Effectiveness in Evaluation of 2.096 points with the following 95% Confidence Interval:

$$0.396 \leq \mu_{\text{Effectiveness-in-Evaluation|Hypertext}} - \mu_{\text{Effectiveness-in-Evaluation|Linear text}} \leq 3.797$$

In Appendix K there is a collection of scatter plot charts of the observed, predicted, and standardized residual values, one for each of the dependent variables, i.e., Chart K1 for Effectiveness in Analysis, Chart K2 for Effectiveness in Synthesis, and Chart K3 for Effectiveness in Evaluation. These charts show no problems with the residuals.

6.3.2 Testing Hypotheses regarding Efficiency

Table 14 includes the tests of Between-Subjects Effects for the dependent variable Efficiency in Evaluation. From this information we can conclude that

Hypothesis H_4, is supported (p-value = 0.010). This means that as expected we do find differences in efficiency between groups working in the two methods (Hypertext, Linear text), when the level of complexity is Evaluation. We can also conclude that the co-variate, Learning Time, does not help us explain this difference (p-value = 0.056).

Source	Type III	df	Mean Sq.	F	p-value	Power
Corrected Model	0.408	2	0.204	4.755	0.011	0.782
Intercept	0.008	1	0.008	0.201	0.655	0.073
Learning Time	0.161	1	0.161	3.754	0.056	0.484
Method	0.235	1	0.235	5.477	0.010	0.640
Error	4.295	100	0.042			
Total	30.637	103				
Corrected Total	4.703	102				

Table 14: Test of Between-Subjects Effects of Efficiency in Evaluation.

Table 15 includes the tests of Between-Subjects Effects for the dependent variable Efficiency in Synthesis. From this information we can conclude that Hypothesis H_5, is not supported (p-value = 0.295). We can also conclude that the co-variate, Learning Time, does not help us in this situation (p-value = 0.713).

Table 16 includes the tests of Between-Subjects Effects for the dependent variable Efficiency in Analysis. From this information we can conclude that Hypothesis H_6, is supported (p-value = 0.304).

Source	Type III	df	Mean Sq.	F	p-value	Power
Corrected Model	0.004	2	0.002	0.221	0.802	0.084
Intercept	0.116	1	0.116	1.249	0.266	0.198
Learning Time	0.001	1	0.001	0.136	0.713	0.065
Method	0.002	1	0.002	0.293	0.295	0.084
Error	9.305	100	0.009			
Total	23.592	103				
Corrected Total	9.346	102				

Table 15: Test of Between-Subjects Effects of Efficiency in Synthesis.

This means that, as expected, we do not find differences in effectiveness between groups working in the two methods (Hypertext, Linear text), when the level of complexity is Analysis. We can also conclude that the co-variate, Learning Time, helps us explain this lack of difference (p-value = 0.001).

Source	Type III	df	Mean Sq.	F	p-value	Power
Corrected Model	0.743	2	0.372	6.889	0.002	0.916
Intercept	0.236	1	0.236	4.368	0.039	0.544
Learning Time	0.673	1	0.673	12.482	0.001	0.938
Method	0.005	1	0.005	1.066	0.304	0.176
Error	5.393	100	0.005			
Total	36.567	103				
Corrected Total	6.136	102				

Table 16: Test of Between-Subjects Effects of Efficiency in Analysis.

6.3.2.1 Parameter Estimates for Efficiency

As in the analysis of the three levels of Effectiveness, we used ANCOVA's mathematical model to express the relationship between the different levels of Efficiency, the co-variate, and the independent variables. We estimated the parameters of these variables using the Deviation Contrasts procedure to calculate the parameters presented in Table 17:

Dependent Variable	Parameter	B	Std. Error	t	p-value
Efficiency Analysis	Intercept	-0.814	0.378	-2.152	0.034*
	Learning Time	0.189	0.054	3.533	0.001*
	Method	0.004	0.046	1.032	0.304
Efficiency Synthesis	Intercept	0.572	0.497	1.150	0.253
	Learning Time	-0.002	0.070	-0.369	0.713
	Method	-0.003	0.060	-0.541	0.295
Efficiency Evaluation	Intercept	-0.199	0.338	-0.590	0.556
	Learning Time	0.009	0.048	1.937	0.056*
	Method	0.010	0.041	2.340	0.010*

Table 17: Parameter Estimates for Efficiency, Method: 1=Hypertext, 0=Linear text.

6.3.2.2 Holm Adjustments for Efficiency

As indicated in the analysis of the adjustment for Effectiveness, we used the Holm simultaneous testing procedure (Holm, 79) to the three levels of Efficiency. In a similar way, we have tested the following three sets of hypotheses:

Test 1: H_o: $\mu_{\text{Efficiency-in-Evaluation} \mid \text{Hypertext}} \leq \mu_{\text{Efficiency-in-Evaluation} \mid \text{Linear text}}$

H_a: $\mu_{\text{Efficiency-in-Evaluation} \mid \text{Hypertext}} > \mu_{\text{Efficiency-in-Evaluation} \mid \text{Linear text}}$

Test 2: H_o: $\mu_{\text{Efficiency-in-Synthesis} \mid \text{Hypertext}} \leq \mu_{\text{Efficiency-in-Synthesis} \mid \text{Linear text}}$

H_a: $\mu_{\text{Efficiency-in-Synthesis} \mid \text{Hypertext}} > \mu_{\text{Efficiency-in-Synthesis} \mid \text{Linear text}}$

Test 3: H_o: $\mu_{\text{Efficiency-in-Analysis} \mid \text{Hypertext}} = \mu_{\text{Efficiency-in-Analysis} \mid \text{Linear text}}$

H_a: $\mu_{\text{Efficiency-in-Analysis} \mid \text{Hypertext}} \neq \mu_{\text{Efficiency-in-Analysis} \mid \text{Linear text}}$

For these test, we chose the same level of significance $\alpha = 0.05$. We ranked the three test according to their p-value. Test 1, with p-value = 0.010 ranks 1, this p-value is compared to $\alpha/3 = 0.05/3 = 0.01667$; H_a is concluded. Then Test 2, with p-value = 0.295 ranks 2. This value is compared to $\alpha/2 = 0.05/2 = 0.025$; H_0 is concluded and the procedure stops. H_0 is concluded for Tests 2 and 3.

Once we have estimated the parameters of MANCOVA's mathematical model for our set of variables and performed the Holm adjustments, we proceed to estimate the 95% confidence intervals for the differences found in Efficiency in Evaluation. Since the difference in performance, as expected, is found to be higher for subjects in the Hypertext group, the difference is calculated as (Hypertext minus Linear text), which gives us a Mean Difference for Effectiveness in Evaluation of 9.56% with the following 95% Confidence Interval:

$$1.45\% \leq \mu_{\text{Efficiency-in-Evaluation}|\text{Hypertext}} - \mu_{\text{Efficiency-in-Evaluation}|\text{Linear text}} \leq 17.7\%$$

Finally, in Appendix K there is a collection of scatter plot charts of the observed, predicted, and standardized residual values, one for each of the dependent variables, i.e., Chart K4 for Efficiency in Analysis, Chart K5 for Efficiency in Synthesis, and

Chart K6 for Efficiency in Evaluation. From these charts, we cannot see any problems with the residuals.

6.4 Data from the Evaluation Form

Data obtained from the evaluation form were analysed in an exploratory manner to determine whether there exist differences between subjects' perceptions of the tools (hypertext, linear text). These data, described in Section 5.5.3 and presented in Table 18, consist of twelve Likert scale keys associated in six factors: PARTICIPATION, CONTENT, TASK, TEST, HARDWARE USE, and GLOBAL SATISFACTION.

FACTOR		SCALE								
		From	1	2	3	4	5	6	7	To
Participation		I am glad I participated								I am not glad I participated
Content	Motivation	I like it								I did not like it
Content	Usefulness	I found it useful								I did not find it useful
Content	Experience	I knew it								I did not know it
Task	Motivation	I like it								I did not like it
Task	Satisfaction	I enjoyed it								I did not enjoy it
Task	User Friendliness	It was friendly								It was unfriendly
Test	Difficulty	It was too easy								It was too difficult
Test	Challenge	Cases were challenging								Cases were not challenging
Test	Motivation	Cases were interesting								Cases were not interesting
Hardware Use		I have use it before								I have never used it before
Global Satisfaction		I had a wonderful experience								I had a terrible experience

Table 18: Scales of Factors from the Evaluation Form

Three of these factors have sub-factors. Factor CONTENT, includes MOTIVATION, USEFULNESS, and EXPERIENCE. Factor TASK includes MOTIVATION, SATISFACTION, and USER FRIENDLINESS. Factor TEST includes GRADE OF DIFFICULTY, CHALLENGE, and MOTIVATION. Their descriptive statistics are presented in Appendix H as Table H1, for the overall sample; Table H2, for the Hypertext Group; Table H3, for the Linear Text Group.

These scales are inverted. A lower number indicates a higher preference and/or a higher degree of satisfaction. In order to see if there were any associations between the factors in the evaluation form with some of the other variables collected, we calculated a correlation matrix of those factors with Total Score, Learning Time, and Total Time. This correlation matrix is presented in Appendix L. Table 19 contains a summary of significant correlations between the evaluation factors and these variables.

We can see that Total Score is correlated with Content (Motivation), Task (Satisfaction), Test (Difficulty and Motivation), and Global Satisfaction. These correlations can be interpreted as those subjects with high scores were more motivated by the content of the learning material, satisfied with their task, found the test less difficult and the cases more motivating, and they were more satisfied globally with their experience.

Both Learning Time and Total Time are correlated with Content (Usefulness), Task (Satisfaction), Test (Challenge and Motivation), and Global Satisfaction. These correlations can be interpreted as indicating that subjects that spent longer in the learning module (and in the evaluation material) found the learning material more useful, were more satisfied with their task, found the test more challenging and the cases more motivating, and in general, were more satisfied with their experience.

Factor			Score	Learning Time	Total Time	Hardware Use
Content	Motivation	r	-0.25			
		Sig.	0.01			
	Usefulness	r		-0.225	-0.217	
		Sig.		0.023	0.029	
Task	Satisfaction	r	-0.2	-0.253	-0.216	
		Sig.	0.05	0.01	0.03	
	User Friendliness	r				0.201
		Sig.				0.043
Test	Difficulty	r	-0.25			
		Sig.	0.01			
	Challenge	r		-0.206	-0.231	
		Sig.		0.038	0.02	
	Motivation	r	-0.22	-0.289	-0.284	
		Sig.	0.03	0.003	0.004	
Global Satisfaction		r	-0.23	-0.203	-0.257	
		Sig.	0.02	0.04	0.009	

Table 19: Significant Correlations of Evaluation Factors ($\alpha<0.05$)

Previous hardware use was only significantly correlated with Task (User Friendliness) which indicates that previous use of similar hardware determined their assessment of user friendliness of the learning material, i.e., those with less experience found it less friendly than those with more experience.

In Appendix M there is a collection of twelve histograms, one for each of the factor and sub-factors of the evaluation form. From these histograms we can see that the distributions are such that, with a sufficiently large sample (n = 102) we can rely on the Central Limit Theorem (Neter et al., 96) to assume sampling distributions of means to be multivariate normal allowing the data to be analysed via MANOVA.

	Value	F-test	df_1	df_2	p-value	power
Hotelling's Trace	0.294	2.179	12	89	0.019	0.922

Table 20: Test of Multivariate Differences by Method

There are a variety of test statistics for evaluating multivariate differences. Table 20 contains Hotelling's trace along with equivalent F test and associated p-value allowing for the testing of the following set of hypotheses:

$$H_0: \mu = \mu_0$$

$$H_a: \mu \neq \mu_0$$

We reject the null hypothesis, H_0, and conclude that there is at least one difference between the dependent variables across the two methods. The power of this test is 92.2% which is a high value.

Once we have found that there is some evidence that there are differences by Method (Hypertext, Linear text), we have to identify these differences. Table 21 contains the tests of Between-Subjects Effects by Method for the twelve factors and sub-factors of the evaluation form. From this Table we can conclude, using a level of significance of α=0.05, that factors CONTENT: USEFULNESS (p-value = 0.012), TASK: SATISFACTION (p-value = 0.017), TASK: USER FRIENDLINESS (p-value = 0.002), and TEST: MOTIVATION (p-value = 0.035) have a significant difference between subjects while all other factors do not.

Since we are considering a series of 12 tests, we used Holm's Simultaneous procedure as described earlier. We chose the same level of significance α = 0.05 for these tests. We ranked them according to their p-value.

Factor		SS	df_1	df_2	F	p-value
Participation		2.282	1	100	1.275	0.262
Content	Motivation	2.189	1	100	1.122	0.292
	Usefulness	11.920	1	100	6.582	0.012
	Experience	2.498	1	100	1.026	0.313
Task	Motivation	8.758	1	100	3.801	0.054
	Satisfaction	10.491	1	100	5.934	0.017
	User Friendliness	13.214	1	100	9.892	0.002*
Test	Difficulty	0.653	1	100	0.559	0.456
	Challenge	0.117	1	100	0.103	0.749
	Motivation	9.318	1	100	4.593	0.035
Hardware Use		0.125	1	100	0.280	0.598
Global Satisfaction		1.583	1	100	1.280	0.261

Table 21: Test of Between-Subjects Effects by Method α=0.05

Factor user friendliness, with p-value =0.002 ranks 1, this p-value is compared to $\alpha/12 = 0.05/12 = 0.0041$; we conclude that there are significant differences in factor user friendliness between the two groups. Then Factor usefulness, with p-value = 0.012 ranks 2. This value is compared to $\alpha/11 = 0.05/11 = 0.0045$; we conclude that there are no significant differences in factor Usefulness and the procedure stops. We conclude that there are no more significant differences in the other factors.

From the factors that do not have a significant difference between groups, we can see that PARTICIPATION, HARDWARE USE, and GLOBAL SATISFACTION are among them. This indicates that subjects in both groups (Hypertext, Linear text) had the same evaluation regarding their participation (mean = 2.13, indicating that they were glad of participating in the experiment), exposure to similar hardware (mean = 1.29, indicating that they have used it before), and were equally satisfied with their participation in the experiment considering all the other factors (mean = 2.36, indicating that "in general they had a wonderful experience.")

Regarding the learning material, the lack of significant differences between subjects in factors CONTENT: MOTIVATION and EXPERIENCE is an indication that subjects in both groups were equally motivated by the learning material (mean = 2.48) and their previous knowledge of it was the same (mean = 3.99).

Subjects were equally motivated by the task. Factor TASK: MOTIVATION measured how much subjects liked working with the learning and testing materials (mean = 2.45). Regarding the Test, the lack of significant differences between subjects in factor TEST: DIFFICULTY indicates that subjects in both groups found the multiple-choice test with the same degree of difficulty (mean = 4.08). Regarding the case studies, subjects in both groups found the cases equally challenging (mean = 2.78).

For the factor with significant differences, user friendliness, we used MANOVA's mathematical model to express the relationship between the two methods. We estimated its parameters using the Deviation Contrasts procedure. These parameters are presented in Table 22.

Once we have estimated the parameters of MANOVA's mathematical model for factor user friendliness, we proceed to estimate its 95% confidence interval.

Factor	Parameter	B	t	p-value
Task:	Intercept	2.220	13.582	0.000
User Friendliness	Method	-0.720	-3.145	0.002

Table 22: Parameter Estimates, Method: 1=Hypertext, 0=Linear Text.

Since this factor has an inverted scale, the difference is calculated as (Linear text minus Hypertext), which give us the following Mean difference, and its respective 95% Confidence Interval:

$$\text{TASK: USER FRIENDLINESS} = 0.720$$

$$0.266 \leq \mu_{\text{TASK: USER FRIENDLINESS} \mid \text{Linear text}} - \mu_{\text{TASK: USER FRIENDLINESS} \mid \text{Hypertext}} \leq 1.174$$

This indicates that subjects in the Hypertext Group found the learning module more user friendly than subjects in the Linear text group.

6.5 Summary

Major research findings presented in this chapter are summarized in Table 23. As expected, the use of Hypertext significantly increased performance in both Effectiveness and Efficiency for the highest level of complexity, i.e., Evaluation. Also, as expected, these differences are not significant for the lowest level of complexity used in this research, i.e., Analysis. At this level it seems that the learning time contributed to explain the differences between subjects. That is explained by the fact that at this level memory plays a major role, more so than in the other two levels. So those subjects with longer learning times, performed better, independently of the method used.

The set of hypotheses regarding Effectiveness and Efficiency in Synthesis were not supported. One possible explanation is that the instrument used at this level, Case 1, was designed to measure subjects' ability to make inferences, based on the model learned, regarding the information of this case study. The outcome of this instrument was the number of correct inferences made. This outcome became a dichotomous variable (either the inference was correct or incorrect) and as such more questions are needed to be able to capture these differences, if they exist. One way to overcome this problem is to modify the instrument and ask for more inferences regarding the information of the case.

Hypotheses		Results
H_1: Users learning from the hypertext version of the learning material will perform more effectively on a case study than users learning from the linear version of the same material, when the emphasis of knowledge is on quantitative and qualitative judgments (i.e., level of complexity is Evaluation).	**H > L**	Supported
H_2: Users learning from the hypertext version of the learning material will perform more effectively on a case study than users learning from the linear version of the same material, when the emphasis is on inferences made about the learning material (i.e., level of complexity is Synthesis).	**H > L**	Not Supported
H_3: Users learning from the hypertext version of the learning material will perform as effectively on a multiple choice test as users learning from the linear version of the same material, when the emphasis is on breaking the model into its constituent parts (i.e., level of complexity is Analysis).	**H = L**	Supported
H_4: Users learning from the hypertext version of the learning material will perform more efficiently on a case study than users learning from the linear version of the same material, when the emphasis of knowledge is on quantitative and qualitative judgments (i.e., level of complexity is Evaluation).	**H > L**	Supported
H_5: Users learning from the hypertext version of the learning material will perform more efficiently on a case study than users learning from the linear version of the same material, when the emphasis is on inferences made about the learning material (i.e., level of complexity is Synthesis).	**H > L**	Not Supported
H_6: Users learning from the hypertext version of the learning material will perform as efficiently on a multiple choice test as users learning from the linear version of the same material, when the emphasis is on breaking the model into its constituent parts (i.e., level of complexity is Analysis).	**H = L**	Supported

Table 23: Summary of Findings

CHAPTER 7. QUALITATIVE DATA ANALYSIS

This chapter contains a report of the analysis of verbal protocols related to performance on problem-solving activities that subjects engaged in while working on the multiple-choice test, and the two case studies. The primary objective of using verbal protocols in this experiment was to verify that subjects actually used the model presented as the learning material while working on the testing material. Verbal protocol data were analysed qualitatively in the form of description and categorization.

First, an introduction to the general issues associated with Verbal Protocol Analysis is presented. The methodology used in this experiment, related to verbal protocols, is described afterwards, including the procedures used to collect, code and analyse the data. A discussion of the implications of this analysis in learning advanced material in a complex and ill-structured domain, conclusions drawn, lessons learned, research findings and contributions end the chapter.

7.1 Verbal Protocol Analysis

Verbal Protocol Analysis is a methodology for collecting and analysing data that relies on verbal reports. According to Ericsson and Simon (84), verbal protocols can be used "*to gain information about the course and mechanisms of cognitive processes of subjects' internal states.*" In Information Systems Research, verbal protocol analysis has been used in various studies related to the understanding general principles of human information processing, the evaluation of human-computer interfaces, and the support of decision-making processes (Todd and Benbasat, 87).

The major concerns of using verbal reports as data are their validity and completeness. Validity can be seriously compromised by careless instructions to verbalise and/or by the researcher's probing, if these issues affect the cognitive process itself. Concurrent verbalization is more prone to interferences than retrospective reports. If researchers ask questions to their subjects during the experiment, that may also be obtrusive to the cognitive process. Completeness refers to whether the verbal data can be comprehensive enough to reflect the course and structure of the underlying cognitive process. Cognitive overload may cause omissions in the reports when subjects are not able to report everything relevant and available in short-term memory (STM) at that time. Information used in parallel tasks may cancel each other out in STM leading to omissions and incomplete reports.

Ericsson and Simon's (84) work evaluated carefully the criticisms of verbal protocol analysis by looking at a large body of think-aloud data. In their analysis, they indicated the conditions under which verbal reports provide an authentic trace of the task-related processes and conditions that result in interference, and thus, in incomplete reports. Their conclusion was that the content of thinking-aloud and immediate retrospective reports, versus those collected at the end of the experiment, were valid. Also, they indicated that the validity and completeness of thinking-aloud and immediate retrospective reports can be enhanced by adopting appropriate methods. So, in general, concurrent verbal reports are more informative than retrospective ones. They also found that thinking-aloud procedures on a wide range of tasks did not appear to change the cognitive processes despite the fact that these procedures needed additional time to complete the reports.

Newell and Simon's (72) Information Processing Theory of Human Cognition was used by Ericsson and Simon (84) as a theoretical framework for their analysis of cognitive processes underlying verbal reports. This theory postulates that a cognitive process can be seen as a sequence of internal states successively transformed by a series of information processes. Each state can be described in terms of small number of information structures, or "chunks," available in the limited storage capacity of STM. An assumption, critical in this framework, is that information must be heeded before it can be verbalised. The optimal time for reporting verbally the thoughts is when these thoughts first enter the subjects' attention as part of their efforts to complete a task. Therefore, concurrent verbal reports are the best way to capture those thoughts.

7.2 Verbalization Methods and Procedures

Thinking-aloud is not entirely alien to everyday life, and almost all subjects have probably had some experience of it before they participate in an experiment. Students at school occasionally are asked to explain orally their solutions to some problems to their fellow students in order to show how they generated their solution. Ericsson and Simon (84) described three different levels at which subjects can verbalise their thought process and content.

7.2.1 Levels of Verbalization

The first level of verbalization is simply the vocalization of oral codes. At this level, there are no intermediate processes, and subjects need to expend no special effort to communicate their thoughts. Still, it is necessary to make a distinction between cases where this process is communicated to themselves, and those where they wish to communicate them to others. The self-directed verbalizations tend to be more casual and idiosyncratic than when they are directed to others. This distinction depends on the subject's interpretation of the instructions, as well as on the verbalization's content.

A second level of verbalization involves description, or rather explication of thought content. At this level, verbalizations do not bring new information into the subject's attention, they only explicate or label information that is not linguistic in nature. This information needs some recoding process. Since the recoding process requires some processing time, subjects verbalizing at this second level are expected to take more time for the task than those not verbalizing. However, Ericsson and Simon (84) hypothesized that such recoding does not change the structure of the process for performing the main task.

A third level of verbalization requires subjects to explain their thought processes or thoughts. An explanation of thoughts, ideas, hypotheses, or motives is not simply a recoding of the information already present in STM. It also requires linking this information to thoughts and information attended to previously. This interpretative process is the main difference between level 2 and 3 verbalizations.

7.2.2 Processes to capture different levels of verbalizations

If, in the normal course of the task at hand, subjects pay attention to the way in which processes are generated in order to bring new information into their STM, their verbalizations should fall within level 2 verbalizations. When this is not the case, i.e., if subjects make intermediate inferences, whether correctly or incorrectly, then their verbalizations are level 3's. This can be directed (intentionally or unintentionally) if subjects are asked to explain their thoughts.

Subjects' protocols are influenced by the *exact wording* of the instructions received. The only common feature to the whole range of techniques used to obtain verbal data is that subjects respond orally to instructions or probes. The crucial issue with the report of verbal protocols is what information is heeded. Two forms of verbal reports are the closest reflection of cognitive processes: *Concurrent verbal reports*, either "talk aloud" or "think aloud" processes, and *retrospective reports*.

In order to apply a more detailed model, Ericsson and Simon (84) made a distinction between the case when subjects utter thoughts that are already encoded in verbal form from the case when subjects re-code verbally and utter thoughts that may be held in memory in some other form (e.g., visually). Using these distinctions, they refer to those utterances as *talking-aloud* and *thinking-aloud* respectively. They also noted that subjects are not aware of the distinction between these two kinds of verbalizations, but differences in instructions may influence them to produce one or the other.

7.3 Verbal Protocols Used in this Study

In order to explore the cognitive process of users' understanding of the learned model in a problem-solving situation, this study included a proportion of subjects in each of the treatment conditions thinking-aloud while working on the multiple-choice test and the two case studies. The focus was the "decisional process of choosing an option" in order to verify if they were using the model presented in the learning module as part of their strategy in answering the testing material. Otherwise, results can be attributed to chance by guessing or by answering in any other idiosyncratic way.

Given the high density of verbal data and the amount of resources needed to analyse them, only five randomly selected subjects in each group were asked to provide them. Verbal protocols were tape recorded, transcribed and matched to the computer-traced data for each of these ten subjects.

Data for one of the subjects in the hypertext group, subject H-888, could not be used. This subject indicated, half way through the testing material, that "*speaking is difficult to me because I'm not [an] English [speaking person], I'm [a] French [speaking one]. So, to express myself while working is difficult. I'm sorry.*" This subject was an exchange student from France taking the System Analysis Course and even though they are required to have a working knowledge of the English language, his verbal protocols were not comparable with the rest of subjects since he was not a member of the same sample. Another subject from the hypertext group, subject H-432, was invited to participate in the verbal protocol group. His verbal reports were used instead of those of subject H-888. Consequently, data from ten subjects were analysed, subjects L-194, L-273, L-570, L-832, and L-943 in the Linear text group and subjects H-123, H-318, H-370, H-432, and H-820 in the Hypertext group.

Verbal protocol data were collected to supplement primarily quantitative measures based on data captured by the tracing program. These quantitative measures contain no indication about why a particular option was chosen, the cognitive process of that decision can only be investigated using process tracing methods, such as verbal protocol analysis (Todd and Benbasat, 87). Without these measures, a study of differences in performance may need the following assumption, which is not explicitly stated in this kind of research: Subjects working on the learning material read it, comprehended it, and were able to transfer it in problem-solving. Similarly, it is assumed that subjects worked with limited time to finish the task, optimized it, and performed as well as they could in order to solve the testing material. This is the reason why cash prize incentives were offered as a way of rewarding subjects' performance.

Verbal protocol data were expected to provide evidence regarding why subjects chose a particular option as the correct one in the multiple-choice test, why they chose a particular statement as their inferences in the questions regarding the first case study, and why they chose a particular statement as their judgment in the questions regarding the second case study. Thus, verbal protocol data may give meaning to the quantitative measures, and help interpret the research findings reported in the previous chapter.

The instructions to verbalize were designed based on the previous expectations. One of the most important issues in their design was to minimize the risk of interfering with the problem-solving activity, since these protocols were only exploratory and the research was mainly concentrated in the measurement of performance. To use level 1 concurrent verbalizations, "talking-aloud protocols," would minimize the risk of interference, but they may generate few complete and informative comments. Level 3 concurrent verbalization, explaining the process of solving the problem while "thinking-aloud" may distract subjects from their primary task, i.e., solving the problem, thus, jeopardizing the validity of the quantitative measurements.

Since the goal was to use a form of verbalization that would not interfere with the cognitive process of task performance yet be informative enough about the normal occurrence of that process, we selected level 2 concurrent verbalizations, "thinking-aloud while working," to balance the concerns previously mentioned. Even though level 2 may not generate as much specific information as level 3, the risk of interfering with the cognitive process was minimized.

Instructions for "thinking-aloud" recommended by Ericsson and Simon (84) were used in this study. These instructions are presented in Appendix N. The goal was to encourage subjects to provide more thorough protocols, discouraging them, at the same time, from explaining their actions and getting into level 3 verbalizations.

Subjects were instructed to read all the information aloud. This represented an additional difficulty, since they had to read frequently. At the time they were reading, there was no other verbal information.

Ericsson and Simon (84) recommend that subjects should be trained with simple tasks if they are not familiar with the procedures of talking-aloud and thinking-aloud verbalizations. Two such training tasks were prepared. They were not used because all subjects providing verbal protocols indicated that they frequently "talk to themselves" while working alone.

In many experiments of thinking, investigators have tried to avoid the difficulty of individual differences in the knowledge held in memory by using "naive" subjects — that is, subjects who did not have any prior experience with the problem domain. Gestalt psychologists were alluding to the difference between naive and experienced subjects in their distinction between productive and reproductive thinking.

In cases where we are able to rule out reproductive, or recognition-based cognitive processes, we can make strong predictions that any information that is recalled must have been heeded in the experimental situation. Ericsson and Simon (84) indicate that the best example they can present of a situation where they can make this prediction is in *learning under experimental control, where a subject changes his/her behaviour into behaviour determined by the experimenter.*

So, basically what Ericsson and Simon (84) claim is a simple hypothesis: The information, heeded during performance of a task, is the information that is reportable, and the information that is reported is information that is heeded. They argue that this hypothesis provides a substantially correct basis for understanding and interpreting verbal reports. Also, they acknowledge that one kind of information that subjects clearly *cannot* report, because it is not available in STM, are the cues that allow them to recognize stimuli. But, the result of the recognition process is heeded and therefore can be reported.

7.4 Data Analysis

Appendix O includes the descriptive statistics for the dependent variables, time variables, and evaluation factors for the Verbal Protocol Group. We included Subject H-888 in that analysis, even though his verbal protocols were discarded. This Appendix also includes the t-test used to test the hypotheses that these two groups (with and without Protocol Analysis) have equal means.

After the basic statistical analysis, data were transcribed and matched to the computer-traced data, all verbal protocols provided as part of reading-aloud were eliminated, leaving only the verbal reports provided while answering the questions of the multiple-choice test and the two case studies. If verbal protocols can be shown to be pertinent to ongoing cognitive processes, they can be used as evidence for the course and nature of these processes. This will also include a consideration of the kinds of inferences about processes that can be drawn from these data. Here we are considering three criteria, regarded as necessary conditions to be satisfied by these protocols if they are to be used to infer underlying cognitive processes.

Relevance Criterion. The protocols should be relevant to the given task. **Consistency Criterion**. The verbalizations, to be pertinent, should be logically consistent with the verbalizations that just preceded them. And, **Memory Criterion**. A subset of the information heeded during task performance will be remembered.

Data that did not comply with these three criteria were omitted from this analysis. Leaving only data that (i) were not read, (ii) were relevant, (iii) were consistent, and (iv) remembered. These data then were coded. The coding of the protocols was based on two aspects of problem-solving strategies: Considering versus Discriminating.

Considering, according to the Webster's Dictionary, indicates that someone has taken all circumstances into account, has considered all things. The dictionary goes into specific details in making a distinction among the verbs **consider** (which basically denotes the directing of the mind to something in order to understand it or to make a decision about it), **study** (which implies more intense concentration of the mind and methodical attention to details), and **weigh** (which suggests a balancing of contradictory information, conflicting opinions or possible eventualities, in reaching a decision).

Discriminating, also according to Webster's, indicates that someone is able to make or perceive fine distinctions. In a similar way, the dictionary goes into specific details in making a distinction among the verbs **discriminate** (which suggests a process of distinguishing minute or subtle differences between similar things), **differentiate** (which suggests the noting or ascertaining of specific differences between things by comparing in detail their distinguishing qualities or features), and **distinguish** (which implies the process of recognising or marking apart from others by special features or characteristic qualities).

Since subjects may refer in their protocols to the language used in the problem, that will be easily recognisable. So they may use the language of the options, the question statement, or the model. Also, since protocols almost always contain information that reveals the subject's control and evaluative processes and goals (Ericsson and Simon, 84), this information will indicate whether subjects are considering (studying, weighing) or discriminating (distinguishing, differentiating). At this point, we are considering these synonyms as one special case. Our goal is to see whether subjects in the hypertext group are able to evaluate the choices by discriminating them based on the model presented in the learning module.

7.4.1 "Distinguishing" Protocols

In order to formalise the encoding of protocols, we used a functional notation of the form **R(x, y, etc.)**, where the R's are relations, and the x, y, etc. are arguments. For the "Distinguishing" Protocols we used the *D* relation, and *a, b, c,* and *d* for the options available in each question. So D(a) will indicate that subjects make a distinction regarding option "a" either by distinguishing it from the rest, discriminating its information, or differentiating it from other options. In other words, the same relation D(x) will be given to any verbalization that indicates that the subject is distinguishing, discriminating, or differentiating.

7.4.2 "Considering" Protocols

For the "Considering" protocols we used relation *C*, so, C(x) indicates that a subject is considering option x, studying it or weighing it. Again, we are giving the same code, C, to any of the synonyms presented previously. When subjects verbalise an option, that will be an indication that they are considering that option. So, previous to any distinguishing protocol there must be a considering one for the same option, but not necessarily after considering an option subjects will discriminate it. They either will consider another option, discriminate it, or choose it as the correct answer. This is based on the previously mentioned criterion of Consistency.

7.4.3 Coding Protocols

By using both of the criteria presented before, we can translate the reports of subjects in both groups. The way to represent an actual choice is by using relation *Ch(x)*, meaning "Subject chose option x." This relation indicates that the processes of distinguishing and considering have ended and subjects made a decision. In encoding vocabulary this means that C(x) either is followed by C(y), D(x), or Ch(x).

Consider the following example. Subject L-194 while working on question four in the multiple choice test reported the following protocols: *"O.K. If the manager's knowledge of the system is minimal, he won't, . . . , he or she won't as readily accept it as well as if their assessment of the system and his support isn't high, he won't be as ready to accept it either, . . . , um . . . if the manager isn't involved in the project, then he won't . . . he or she won't be readily . . . ready to accept it either. It's manager job characteristics, then."*

In Question 4, the options are (*a*) manager assessment of system and support, (*b*) manager knowledge of system, (*c*) manager involvement, and (*d*) manager job characteristics. By replacing the options by their letter we have: *"O.K. If b is minimal, he won't, . . . , he or she won't as readily accept it as well as if a isn't high, he won't be as ready to accept it either, . . . , um . . . if no c, then he won't . . . he or she won't be readily . . . ready to accept it either. It's d, then."* Finally, this report can be coded as:

"C(b) D(b) C(a) D(a) C(c) D(c) C(d) D(d) C(d) Ch(d)"

Let's compare this coded protocol to the one given by subject H-370 working on the same question. *"Manager assessment of system and support does not have a direct influence . . . acceptance . . . mm . . . manager assessment of system and support . . . of system and support does . . . manager knowledge of systems does . . . manager involvement . . . manager involvement does not."* Again, after substituting the options it becomes: *"a does not have a direct influence . . . acceptance . . . mm . . . a . . . a does . . . b does . . . c . . . c does not."* Finally, this report was coded as:

"C(a) D(a) C(a) C(a) D(a) C(b) D(b) C(c) C(c) D(c) C(c) Ch(c)"

Note that in the previous example there is no indication of considering or discriminating option d. Subject H-370 chose option c as soon as he recognised its special characteristic qualities according to the model learned. He went from a to b to c and stopped there, when he recognised it as the correct answer. Also note that in the case of subject L-194, he did not start with option a, he went from option b to a to c to d and finally chose d as his answer.

Let's consider another example. Subject L-570 working on question five, with options (*a*) manager decision style, (*b*) top management support, (*c*) manager involvement, and (*d*) manager belief in system concept. "*Top management support . . . only one of the following variables has a direct influence . . . um . . . manager decision . . . getting confused between direct influence and non direct influence . . . mm . . . top management support . . . manager belief . . . direct influence . . . manager acceptance . . . manager decision style . . . that's it.*" This report was coded as:

"C(b) D(b) C(a) D(a) C(b) C(d) D(d) C(a) Ch(a)"

Once again, note that this subject did not consider option c, started with option b, then moved to a, then went back to b, move to d, then to a, and chose a as her answer. She also indicated confusion regarding the relationship of influence between manager acceptance and the options which was presented in the learning material. This confusion is an indication of lacking a criterion to make an assessment regarding what does it mean to have a dimension or variable directly (or indirectly) influencing another dimension. So, she was unable to distinguish the options according to the relevant criterion.

Unfortunately after the codes were assigned, we noted that there was a large proportion of incomplete verbalizations which resulted in inconsistent codes. These codes were discarded, leaving a sample that was no longer useful. What follows are some examples of incomplete protocols and inconsistent codes.

Consider subject L-832 working on the second case study. He provided the following verbalizations: "*Er . . . discussion of assessment . . . er, there hasn't been any sort of mention about the quality of, er, ESOP . . . er . . . I'm unclear whether or not his assessment, or how his assessment was . . . what his assessment was . . . so I'll have to say there's not enough information to determine.*" Which translate into the following code:

"C(a) D(a) C(d) Ch(d)"

Also, subject H-820's verbalizations while working on the same question were "*knows that the system has great support but is not quite satisfied . . . likes ESOP but knows he'll have to work hard to have . . . joined the company mainly due to ESOP. He knew that the organization was ready for such a system . . . there is not enough information to determine assessment . . . review . . . light . . . understand characteristics . . . the natural habitat . . . O.K. . . . ESOP . . . decide . . . validate . . . O.K., O.K., O.K. questions . . . likes [inaudible] . . . he likes it . . .*" And its code

"C(a) C(b) C(c) C(d) C(b) Ch(b)"

Finally, consider subject L-943's report while working on the second case study. "*It's most likely Zerimar needed to show a tool for the impact of ESOP to all members because . . . hm . . . um . . . I just clicked, I think, four questions the same. I* *thought they were the same thing - I thought it was [inaudible]. I think this is the first question for the . . . for this case. I clicked four times the bottom. Well, we can just . . . the first question I thought I answered right. And at 12 I'm just going to start . . . put everything . . . between . . .* " Thus, his report for question nine was eliminated, and no reports were provided for questions ten and eleven.

7.5 Discussion

The primary objective of using verbal protocols in this experiment was to verify that subjects actually used the model presented as the learning material while working on the testing material. Then, these verbal reports allowed us to look for a way to verify that subjects actually used the model presented in the learning material as part of their problem-solving strategies. That is clearly observed in the protocols. Some of the linear model subjects reported confusion regarding the kind of interactions between dimensions and variables, or among dimensions. No subjects in the hypertext group reported a similar problem. For them it was quite straightforward since they had links indicating these interactions.

Even though we used concurrent verbal reports as a way to capture subjects' thoughts, which is considered the preferred method, we ended up having a problem of incomplete reports. Since, completeness refers to whether the verbal data can be comprehensive enough to reflect the course and structure of the underlying cognitive process, we consider this a major setback in our objective of analysing differences between groups. Cognitive overload may have caused omissions in the reports, but somehow subjects were not able to report everything relevant and available in their short-term memory (STM) at that time. As Ericsson and Simon (84) suggested, information used in parallel tasks may have cancelled each other out in STM leading to omissions and incomplete reports. At the end, 43.33% of the codes obtained by these subjects working on the multiple-choice test (Analysis), 30.00% of the codes obtained from the first case study (Synthesis), and 40.00% of the codes obtained from the second case study (Evaluation) were eliminated.

This analysis was intended to measure the implications of using a hypertext-based learning aid to learn advanced material in a complex and ill-structured domain. At this point we are unable to provide a reliable measure of those implications, if any.

By observing these reports we can see that subjects were referring to the model learned while weighing the alternatives before they chose one of them as the correct answer. They were also using the language provided in the learning module, the questions, and the options, i.e., the language of Lucas, Ginzberg and Shultz' model. This may represent a way to indicate that subjects engaged in the learning process. Whether there is a difference between the experimental and the control groups in the way they used this model in a problem-solving situation is still to be seen. At this point there is no enough evidence to draw any conclusion in that direction.

By engaging subjects into a thinking-aloud process we were looking for a way to validate their observed differences in accuracy and speed in a problem-solving task. As Ericsson and Simon (84) indicate, "*subjects' protocols are influenced by the exact wording of the instructions received.*" We have learned that our instructions, as they were presented to our subjects, called for verbal reports closer to level 1 — talking-aloud — than to level 2 — thinking-aloud. Since the only common feature to the whole range of techniques used to obtain verbal data is that subjects respond orally to instructions or probes, this indicates that we captured the kind of verbal reports that we requested, not necessarily the kind we were aiming for in this experiment.

Even though these reports were not categorised and grouped, we have some evidence that subjects working on the linear text model were unable to learn some aspects of the model. Aspects that may have an impact on how they performed. This difference in performance can be explained by the way information was accessed in each of the two groups. In the linear text module, information was presented sequentially, while in the hypertext module it was linked. These links denoted the type of association that exists between dimensions and variables, i.e., direct or indirect. So, when subjects were asked about these differences, members of the linear text group reported confusion. Thus, empirical results can be interpreted within these differences.

Two aspects to consider in future research using verbal protocols include the need for more clearly stated instructions and the training tasks suggested by Ericsson and Simon (84). With clearer instructions, subjects will be able to provide more complete protocols while engaged in the problem-solving activities. With the training tasks, the researcher will be able to indicate to subjects to which extent their protocols are relevant, consistent and remembered.

If time and resources are available, a larger sample size will be more useful. With a larger sample, even in the case of having incomplete reports, the proportion of useful ones to the ones discarded, still will be large enough for analysis, although the amount of work may escalate rapidly.

CHAPTER 8. CONCLUSIONS

This chapter concludes this dissertation. It includes a discussion of the major research contributions, implications of these findings, limitations, and directions for future research. The major purpose of this study was to make an assessment regarding the enhancement of hypertext-based learning aids in the acquisition of advanced knowledge. This enhancement was evaluated by looking at the usefulness of hypertext as a knowledge acquisition tool, when implemented according to the Cognitive Flexibility Theory. A hypertext-based learning aid and a linear text learning aid were developed and tested in order to conduct this research. Subjects were randomly assigned to these treatments, i.e., hypertext and linear text. Subjects in the hypertext group were designated as the experimental group and those in the linear text group were designated as the control group.

8.1 Research Contributions and Implications

8.1.1 Summary of Findings

Two sets of three hypotheses were formulated and tested. The first set is concerned with effectiveness of performance in problem-solving, while the second is concerned with efficiency. In both cases, hypotheses were tested at the $\alpha=0.05$ level of significance and adjusted for multiple comparisons using Holm's Multiple Comparison Procedure (Holm, 79). Discussion of these hypotheses and the findings is presented first for the hypotheses dealing with effectiveness, while the hypotheses dealing with efficiency are presented later.

8.1.2 Hypotheses regarding effectiveness

> Hypothesis 1: Users learning from the hypertext version of the learning material will perform more effectively on a case study than users learning from the linear version of the same material, when the emphasis of knowledge is on quantitative and qualitative judgments (i.e., level of complexity is Evaluation).

Significant differences were detected in effectiveness between the hypertext and linear text groups. Those subjects in the hypertext group were able to provide more accurate judgments than those in the linear text group while working in a case study. This hypothesis was supported. We concluded that the hypertext version of the learning material was a significant factor in performance at this level of complexity.

> *Hypothesis 2*: Users learning from the hypertext version of the learning material will perform more effectively on a case study than users learning from the linear version of the same material, when the emphasis is on inferences made about the learning material (i.e., level of complexity is Synthesis).

No significant differences in effectiveness were found between the hypertext and the linear text groups while making inferences working in a case study. Therefore, this hypothesis was not confirmed. We concluded that the version of the learning material was not a significant factor in performance at this level of complexity.

> Hypothesis 3: Users learning from the hypertext version of the learning material will perform as effectively on a multiple choice test as users learning from the linear version of the same material, when the emphasis is on breaking the model into its constituent parts (i.e., level of complexity is Analysis).

No significant differences in effectiveness were found between subjects of the hypertext and linear text groups while working in a multiple-choice test. Therefore, this hypothesis was supported. We conclude that at this level of complexity the learning material was not a significant factor in performance.

8.1.3 Hypotheses regarding efficiency

> *Hypothesis 4*: Users learning from the hypertext version of the learning material will perform more efficiently on a case study than users learning from the linear version of the same material, when the emphasis of knowledge is on quantitative and qualitative judgments (i.e., level of complexity is Evaluation).

Significant differences were detected in efficiency between the hypertext and linear text groups. Those subjects in the hypertext group were able to provide more efficient judgments than those in the linear text group while working in a case study. This hypothesis was supported. We concluded that the hypertext version of the learning material was a significant factor in performance at this level of complexity.

> *Hypothesis 5*: Users learning from the hypertext version of the learning material will perform more efficiently on a case study than users learning from the linear version of the same material, when the emphasis is on inferences made about the learning material (i.e., level of complexity is Synthesis).

No significant differences in efficiency were found between the hypertext and the linear text groups while making inferences working in a case study. Therefore, this hypothesis was not confirmed. We concluded that the version of the learning material was not a significant factor in performance at this level of complexity.

> *Hypothesis 6*: Users learning from the hypertext version of the learning material will perform as efficiently on a multiple choice test as users learning from the linear version of the same material, when the emphasis is on breaking the model into its constituent parts (i.e., level of complexity is Analysis).

No significant differences in efficiency were found between subjects of the hypertext and linear text groups while working in a multiple-choice test. Therefore, this hypothesis was supported. We conclude that at this level of complexity the learning material was not a significant factor in performance.

8.1.4 Implications of the Findings

The previously presented findings suggest that subjects' performance is more effective and efficient by using a hypertext-based learning aid at a higher level of complexity than at lower levels. This can be interpreted as, hypertext-based learning aids significantly improve problem-solving performance of subjects at higher levels of complexity. In other words, there exists cognitive fit between the level of complexity and hypertext-based learning aids.

This implies that educators or trainers can benefit from hypertext by using this kind of learning aid to address complex and ill-structured learning materials. To insure that resources are better allocated and not wasted, educators and trainers could insist that learning material that is complex and ill-structured may be implemented in hypertext. Also, they may insist that only that kind of material be implemented in hypertext, since there is a direct improvement on performance.

A major contribution of this research is the methodology used to assess the cognitive fit of more complex pairs of tools-tasks than those reported previously in the literature. The research design and the experimental materials were based on the "match" between the technology and the task, under the cognitive fit paradigm. It is within this theoretical framework that the effects of hypertext and level of complexity were analysed and hypothesised. Intervening problem-solving processes were linked to verbal protocols to help explain the effects of the independent variable on the dependent variables.

The material was designed within the philosophy of constructivism. One of the assumptions that works well in this model is that meaning varies with how the individual creates meaning from his or her experiences. Since no subjects in the hypertext module chose to traverse the information sequentially, each one of them had a different learning experience. Each of them created her or his own construction of the model. This did not mean that some were right and others wrong; the model learned is the same, their understanding of it was different.

An important feature of the hypertext learning material used is that it allowed learners to selectively visit each dimension and the factors/dimensions that have a direct/indirect influence on them. This option provided a way to understand the patterns of association (by means of the highlighted information) which revealed the conceptual information of the model. This was observed on some of the verbal reports provided by subjects in the verbal protocol group. We can say that just by exploring, in a constructivist manner, subjects in the hypertext module were actually learning.

8.2 Limitations

One of the major limitations of this research is the method chosen for this study. We used a laboratory experiment, emphasising internal validity, but weak external validity. Despite considerable efforts to minimize the limitation, external validity suffers from the artificiality of a laboratory setting and data collection procedures. Thus, findings are less likely to be generalised across settings (Cook and Campbell, 79).

On one hand, we had an artificial setting, on the other, we used a complex and well-supported implementation model. The setting was artificial mainly because subjects worked in a different environment than they regularly use, i.e., a classroom. Otherwise, we can claim that the learning environment is similar to any other computer-based learning material. The content, though, is an actual implementation model. This model was implemented directly from the book it was reported without any simplification.

Subjects were undergraduate students, majoring in MIS, enrolled in a System Analysis Course. For them, the learning material is similar to the one included in their syllabus for the course, even though this particular model is not included. Therefore, it may still be appropriate to generalise the findings to similar populations, i.e., MIS majoring undergraduate students.

Differences in performance were measured by having subjects work on the same problem-solving tasks. The experiment took place in a two-week period. It is possible that subjects commented among themselves the kind of problem-solving tasks used. But, since they were unaware of exactly what options were correct, it would not help them much to know in advance the kind of tasks they would be evaluated.

The cases used were validated as a reliable tool that provided similar results when used by different subjects. They were not validated against other similar cases. This cross validation is needed, in order to generalise these conclusions.

One important limitation is the fact that complex models are difficult to implement. The information necessary to show differences that exist, mainly in perception, can rapidly escalate. Sometimes it may not be feasible to attempt its implementation; Other times, it may be irrelevant to do so. One of the most important learning experiences, derived from this research, is the acknowledgement that finding a workable balance between feasibility and relevancy should be the main goal when designing instructional material for hypertext.

8.3 Directions for Future Research

Hypertext is not the solution for every problem in learning. There are many aspects of learning where hypertext does not have a meaningful contribution. That was clearly shown in this research. Still, we need some criteria to determine which aspects of the learning process are better served with hypertext. It seems that those areas where information is not linear, or where by imposing linearity the information is not oversimplified, are excellent choices.

The design of the new generation of hypertext learning aids should foster learning by discovery in a constructivist way, bring some kind of structure to ill-structured material, and allow material to be revisited at different times, in different contexts. In cases where the amount of information is overwhelming, that information should be broken into modules and meaningful links should allow the traversal of each of them from different parts of the system. When necessary, the system should provide "guided tours" to the information. One aspect to investigate on these guided tours, is whether they promote motivation to learn.

This study can be extended in at least two ways. First, it may be worth investigating the application of hypertext-based learning aids to other tasks dealing with knowledge acquisition. Second, a similar assessment of Cognitive Fit can be performed using other tools (multimedia, World Wide Web, Internet, Intranet, etc.) for similar tasks (learning). We can use Leidner and Jarvenpaa's (95) theoretical information technology fit with other theories of learning, to match the task-tool pairs.

With the advance in Information Technology applications, we expect to find more applications of these technologies in schools, universities, and organizations. If that is the case, it will be possible to use those tools instead of having to create them specially for research studies, increasing the external validity of research findings.

In conclusion, this research is the beginning of what is anticipated to be a rich stream of research based upon measuring the effects of emerging technologies on training and learning. It seems that this kind of research in not only needed, it is also urgent. Emerging technologies are having more and more an impact on the way we educate and train in particular, and do businesses in general.

BIBLIOGRAPHY

ACM, (1988) **Hypertext 87: Special Issue on Hypertext**. *Communications of the ACM*, 31, 7, July.

Adams, J. A., (1993) **Applications, Implications**. *IEEE Spectrum*, March, 24-31.

Alexander, J. W. and Randolph, W. A. (1985) **The fit between technology and structure as a predictor of performance in nursing subunits**. *Academy of Management Journal*, 28, 844-859.

Anderson, J. R., (1983) **The Architecture of Cognition**. Cambridge, MA: Harvard University Press.

Bandura, A., (1977) **Social Learning Theory**. Englewood Cliffs, NJ: Prentice Hall

Bettman, J. R. and Kakkar, P., (1977) **Effects of information presentation format on consumer information acquisition strategies**. *Journal of Consumer Research*, 3, 233-240.

Bettman, J. R. and Zins, M., (1979) **Information Format and Choice Task in Decision-Making**. *Journal of Consumer Research*, 6, 141-153.

Bloom, B. S. (Ed.) (1956) **Taxonomy of Educational Objectives: The classification of Educational Goals**. Handbook I: Cognitive Domain. New York, NY: D. McKay Co.

Bloor, D., (1983) **Wittgenstein and the social construction of knowledge**. New York, NY: Columbia University Press.

Brown, J. S., (1985) **Process versus Product: a perspective on tools for communal and informal electronic learning**, *Journal of Educational Computing Research*. 1, 179-202.

Brown, J. S., Collins, A., and Duguid, P., (1989) **Situated Cognition and the Culture of Learning**. *Educational Researcher*, 18, 32-42.

Card, S. K., Moran, T. P., and Newell, A., (1983) **The Psychology of Human Computer Interaction**. Hillsdale, NJ: Lawrence Erlbaum Associates.

Catano, J., (1989) **Poetry and Computers: Experimenting with Comunaltext.** *Computers and the Humanities*, 13, 269-275.

Chandler, D. (1990) **Hyperexpectations.** *UNIX Review*, 11, 2, 5.

Cognition and Technology Group at Vanderbilt University, (1992) **Technology and the Design of Generative Learning Environments**. In Duffy, T., and Jonassen, D. H. (Eds.) *Constructivism and the Technology of Instruction: A Conversation.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J., (1988) **Statistical Power Analysis for the Behavioral Sciences**. 2nd. Edition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Collins, A., Brown, J. S., and Newman, S. E., (1989) **Cognitive Apprenticeship: Teaching the crafts of reading, writing, and mathematics**. In Resnick, L.B. (Ed.) *Knowing, Learning, and Instruction: Essays in Honor of Robert Glaser*. Hillsdale, NJ: Lawrence Erlbaum Associates, 453-494.

Cook, T. D., and Campbell, D. T., (1979) **Quasi-Experimentation: Design and Analysis Issues for Field Settings**. Boston, MA: Houghton Mifflin Company.

Dionne, D. L., and Krull, R., (1992) **The Effects of Module Size and User-Practice in Online Tutorials**. In Society for Technical Communication's *Proceedings of the 39th Annual Conference*, Arlington, VA, 669-672.

Drazin, R. and Van de Ven, A. H., (1985) **An examination of alternative forms of fit in Contingency Theory**. *Administrative Science Quarterly*, 30, 514-539.

Duffy, T. M. and Jonassen, D. H., (1992) **Constructivism and the Technology of Instruction: A Conversation**. Hillsdale, NJ: Lawrence Erlbaum Associates.

Duffy, T. M., and Knuth, R. A., (1990), **Hypermedia and Instruction: where is the match?** In Jonassen, D. H., and Mandl, H., (Eds.) *Designing Hypermedia for Learning*. NATO ASI Series F, Vol. 67. Berlin: Springer Verlag, 199-226.

Ebel, R. L. and Frisbie, D. A., (1986) **Essentials of Educational Measurement**. 4th Edition. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Eberts, R., (1988) **Computer-based Instruction**. In *Handbook of Human-Computer*

Interaction. Amsterdam: North-Holland, 599-627.

Eberts, R. and Brock, J. F., (1984) **Computer Applications to Instruction**. *Human Factors Review*, 239-284.

Eckert, P., (1989) **Jocks and Burnouts.** Philadelphia, PA: Temple University Press.

Ericsson, K. A., and Simon, H. A., (1984) **Protocol Analysis: Verbal Reports as Data**. Cambridge, MA: The MIT Press.

Ess, C., (1991), **The pedagogy of computing: Hypermedia in the classroom**. In *Proceedings of the 3rd ACM Conference on Hypertext (Hypertext '91)*. New York, NY: ACM Press, 277-290.

Francis, L., (1992) **W(h)ither Go Tutorials**. In Society for Technical Communications' *Proceedings of the 39th Annual Conference*, Arlington, VA, 412-415.

Fosnot, C., (1989) **Enquiring Teachers, Enquiring Learners: A constructivists approach for teaching.** New York, NY: Teachers College Press.

Gagné, R. M., (1985) **The Conditions of Learning and Theory of Instruction**. 4th Edition, New York, NY: Holt, Rinehart and Winston.

Gery, G., (1987) **Making CBT Happen: Prescriptions for successful Implementation of Computer-Based Training in your Organization**. Boston, MA: Weingarten Publications.

Gordon, S., Gustavbel, J., and Hankey, J., (1988) **The Effects of Hypertext on Reader Knowledge Representation**. *Proceedings of the Human Factors Society, 32nd Annual Meeting*, 296-300.

Hammond, N., (1989) **Hypermedia and Learning: who guides whom?** In Maurer, H., (Ed.) *Computer Assisted Learning: Proceedings of the 2nd International Conference, ICCAL '89*. Berlin: Springer-Verlag, 167-181.

Holm, S., (1979) **A Simple Sequentially Rejective Multiple Test Procedure**. *Scandinavian Journal of Statistics*, 6, 65-70.

Jacobson, M. J., (1994) **Issues in Hypertext and Hypermedia Research: Toward a Framework for Linking Theory-to-Design.** *Journal of Educational Multimedia and Hypermedia*, 3, 2, 141-154.

Jones, R. A., and Spiro, R., (1992) **Imagined Conversations: The relevance of Hypertext, Pragmatism, and Cognitive Flexibility Theory to the interpretation of "Classic Texts" in Intellectual History**. In Lucarella, D. et al. (Eds.) *Proceedings of the ACM Conference on Hypertext: ECHT '92*. Salem, MA: ACM Publications, 141-148.

Jonassen, D. (1993) **The trouble with Learning Environments** *Educational Technology* 23, 1, January, 35-37.

Jonassen, D. and Grabinger, R., (1990) **Problems and Issues in Designing Hyper-text/Hypermedia for Learning**. In Jonassen, D. and Mandl, H., (Eds.) *Designing Hypermedia for Learning*. NATO ASI Series, Series F: Computer and Systems

Sciences, Vol. 67, Berlin: Springer-Verlag.

Jonassen, D. and Mandl, H., (1990) **Designing Hypermedia for Learning.** NATO ASI Series, Series F: Computer and Systems Sciences, Vol. 67, Berlin: Springer-Verlag.

Joyce, W., Slocum, J. and Von Glinow, M., (1982) **Person-situation interaction: Competing models of fit**. *Journal of Occupational Behavior*, 3, 265-280.

Khalifa, M. and Ramírez, A., (1992) **A Hypercourseware for MIS.** *Proceedings of the 7th Annual Conference of the IAIM*. Dallas, Tx. December, 431-441.

Kearsley, G., (1985) **Embedded Training: The new look of Computer-based Instruction**. *Machine Mediated Learning*, Vol. 1, No. 3, 279-285.

Lakatos, I., (1970) **Falsification and the Methodology of Scientific Research Programmes**. In Lakatos, I., and Musgrave, A., (Eds.) *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press.

Lave, J., (1988) **Cognition in Practice.** Cambridge, MA: Cambridge University Press.

Leidner, D. and Jarvenpaa, S., (1995) **The use of Information Technology to Enhance Management School Education: A Theoretical View**. *MIS Quarterly*, 19, 3, September, 265-291.

Legget, J. J., Schnase, J. L., and Kacmar, C. J., (1990) **Hypertext for Learning.** In Jonassen, D. and Mandl, H. (Eds.) *Designing Hypermedia for Learning*. Berlin:

Springer-Verlag, Chapter 2, 27-37.

Locatis, C., Letourneau, G., and Banvard, R., (1989) **Hypermedia and Instruction.** *Educational Technology Research and Development*. 37, 4, 65-77.

Lucas, Jr., H. C., Ginzberg, M. J., and Schultz, R. L. (1990) **Information Systems Implementation: Testing a structural model**. Norwood, NJ: Ablex Publishing Co.

Madigan, D., Chapman, C. R., Gavrin, J., Villumsen, O., and Boose, J., (1994) **Repertory Hypergrids: an application to clinical practice guidelines**. In *Proceedings of the ACM European Conference on Hypermedia Technology ECHT '94*. Glasgow: ACM Press, 117-125.

Maturana, H. and Varela, F., (1987) **The tree of knowledge: The biological roots of human understanding**. Boston, MA: Shambhala.

Michalak, S., and Coney, M., (1993) **Hypertext and the Author/Reader Dialogue**. In *Proceedings of the 5th ACM Conference on Hypertext (Hypertext '93)*. New York, NY: ACM Press, 174-182.

Nelson, W. A., and Palumbo, D. B., (1992) **Learning, Instruction and Hypermedia.** *Journal of Educational Multimedia and Hypermedia*, 1, 3, 287-299.

Neter, J., Kutner, M., Nachtshein, C., and Wasserman, W. (1996) **Applied Linear Statistical Models**, 4th Edition. Chicago, IL: IRWIN.

Newell, A. and Simon, H. (1972) **Human Problem Solving** Englewood Cliffs, NJ: Prentice Hall.

Norman, D. A., (1981) **Twelve Issues for Cognitive Science.** In Norman, D. A. (Ed.) *Perspectives on Cognitive Science*. Norwood, NJ: Ablex Publishing Co., Chapter 11, 265-295.

Price, R. V., (1991) **Computer-Aided Instruction: A guide for authors**. Pacific Grove, CA: Brooks/Cole.

Radlinski, B., and McKeendre, J., (1992) **Grace Meets the "Real World:" Tutoring COBOL as a second language**. In ACM's *Human Factors in Computer Systems: CHI'92 Proceedings*, New York, NY: ACM Press, 343-350.

Rasch, G., (1960) **Probabilistic Models for some Intelligence and Attainment Tests**. Chicago, IL: University of Chicago Press.

Rasch, G., (1961) **On general laws and the meaning of measurement in psychology**. In *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, CA: University of California Press. 321-333.

Rushby, N., (1987) **From Trigger Video to Videodisc: A case study in Interpersonal Skills**. In Laurillard, D. (Ed.) *Interactive Media: Working Methods and Practical Applications*. Chicester, England: Ellis Horwood, Ltd., 116-131.

Salomon, G., Perkins, D. N., and Globerson, T., (1991) **Partners in cognition: Extending**

human intelligence with intelligent technologies. *Educational Researcher*. 20, 3, 2-9.

Scardamalia, M., and Bereiter, C., (1991) **Higher Levels of Agency for Children in Knowledge Building: A Challenge for the Design of New Knowledge Media.** *Journal of the Learning Sciences,* 1, 1, 37-68.

Schein, E. H., (1992) **The role of the CEO in the Management of Change: The case of Information Technology**. In Kochan, T. A. and Useem, M. (Eds.) *Transforming Organizations*. New York, NY: Oxford University Press, 80-95.

Scribner, S., (1984) **Studying working intelligence**. In Rogoff, B. and Lave, J. (Eds.) *Everyday Cognition: Its development in Social Context.* Cambridge, MA: Harvard University Press.

Skinner, B. F., (1953) **Science and Human Behavior**. New York, NY: Free Press.

Skinner, B. F., (1969) **Contingencies of Reinforcement**. East Norwalk, CT: Appleton-Century-Crofts.

Skinner, B. F., (1971) **Beyond Freedom and Dignity**. New York, NY: Bantam Books.

Simon, H. A. and Hayes, J. R., (1976) **The understanding process: Problem isomorphs**. *Cognitive Psychology*, 8, 165-190.

Spiro, R. J., Coulson, R. L., Feltovich, P. J., and Anderson, D. K., (1988) **Cognitive**

Flexibility Theory: Advanced Knowledge Acquisition in Ill-structured Domains. In *Proceedings of the 10th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Lawrence Erlbaum Associates, 375-383.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., and Coulson, R. L., (1992) **Cognitive Flexibility, Constructivism, and Hypertext: Random Access Instruction.** In Duffy, T. M., and Jonassen, D. H., (Eds.) *Constructivism and the Technology of Instruction: A Conversation.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Spiro, R. J., and Jehng, J., (1990) **Cognitive Flexibility and Hypertext: Theory and Technology for the Nonlinear and Multidimensional Traversal of Complex Subject Matter.** In Nix, D. and Spiro, R. J. (Eds.) *Cognition, Education and Multimedia: Exploring Ideas in High Technology.* Hillsdale, NJ: Lawrence Erlbaum Associates, 163-205.

Spiro, R. J., Vispoel, W. P., Schmitz, J. G., Samarapungavan, A., and Boerger, A. E. (1987) **Knowledge Acquisition for Application: Cognitive Flexibility and Transfer in Complex Content Domains.** Technical Report No. 409, University of Illinois at Urbana-Champaign.

Tergan, S.O., (1997) **Misleading Theoretical Assumptions in Hypertext/Hypermedia Research,** *Journal of Educational Multimedia and Hypermedia,* 6, 3/4, 257-283.

Todd, P., and Benbasat, I., (1987) **Process Tracing Methods in Decision Support Systems Research: Exploring the Black Box.** *MIS Quarterly,* 11, 4, December, 492-512.

Van Dam, A., (1988) **Hypertext'87 Keynote address.** *Communications of the ACM*, 31, 7, July, 887-895.

Van de Ven, A.H. and Drazin, R., (1985) **The concept of fit in contingency theory.** In Cummings and Staw (eds.) *Research in Organizational Behavior*, 7, 333-365.

Varela, F., (1989) **Connaître: Les Sciences Cognitives, tendandes et perspectives**. Paris: Seuil.

Venkatraman, N., (1989) **The concept of fit in strategy research: Toward verbal and statistical correspondence**. *Academy of Management Review*, 14, 3, 423-444.

Venkatraman, N. and Camillus, J. C., (1984) **Exploring the concepts of "fit" in strategic management**. *Academy of Management Review*, 11, 71-87.

Vessey, I., (1991) **Cognitive Fit: A Theory-based Analysis of the Graphs versus Tables Literature**. *Decision Sciences*, 22, 219-240.

Vessey, I. and Galletta, D., (1991) **Cognitive Fit: An empirical Study of Information Acquisition**. *Information Systems Research*, 2, 1, 63-84.

Willis, P., (1981) **Learning to labour**. New York, NY: Columbia University Press.

Winograd, T., and Flores, F., (1986) **Understanding Computers and Cognition: A New Foundation for Design**. Norwood, NJ: Ablex Corporation.

Wittgenstein, L., (1953) **Philosophical Investigations**. New York, NY: MacMillan.

Wright, B. D., and Stone, M. H., (1979) **Best Test Design: The Rasch Measurement**. Chicago, IL: MESA Press.

Yankelovich, N., Haan, B., Meyrowitz, N., and Drucker, S., (1988) **Intermedia: The construction of a seamless information environment.** *IEEE Computer,* 21, 1, 81-96.

Zuboff, S., (1988) **In the Age of the Smart Machine**. New York, NY: Basic Books.

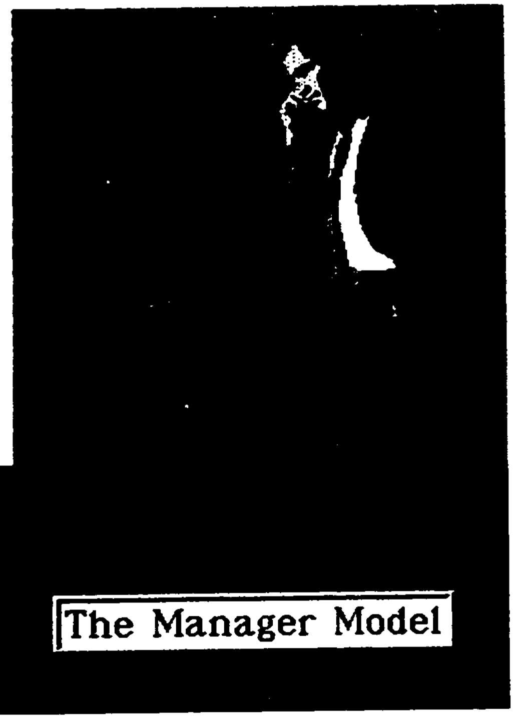
Appendix A

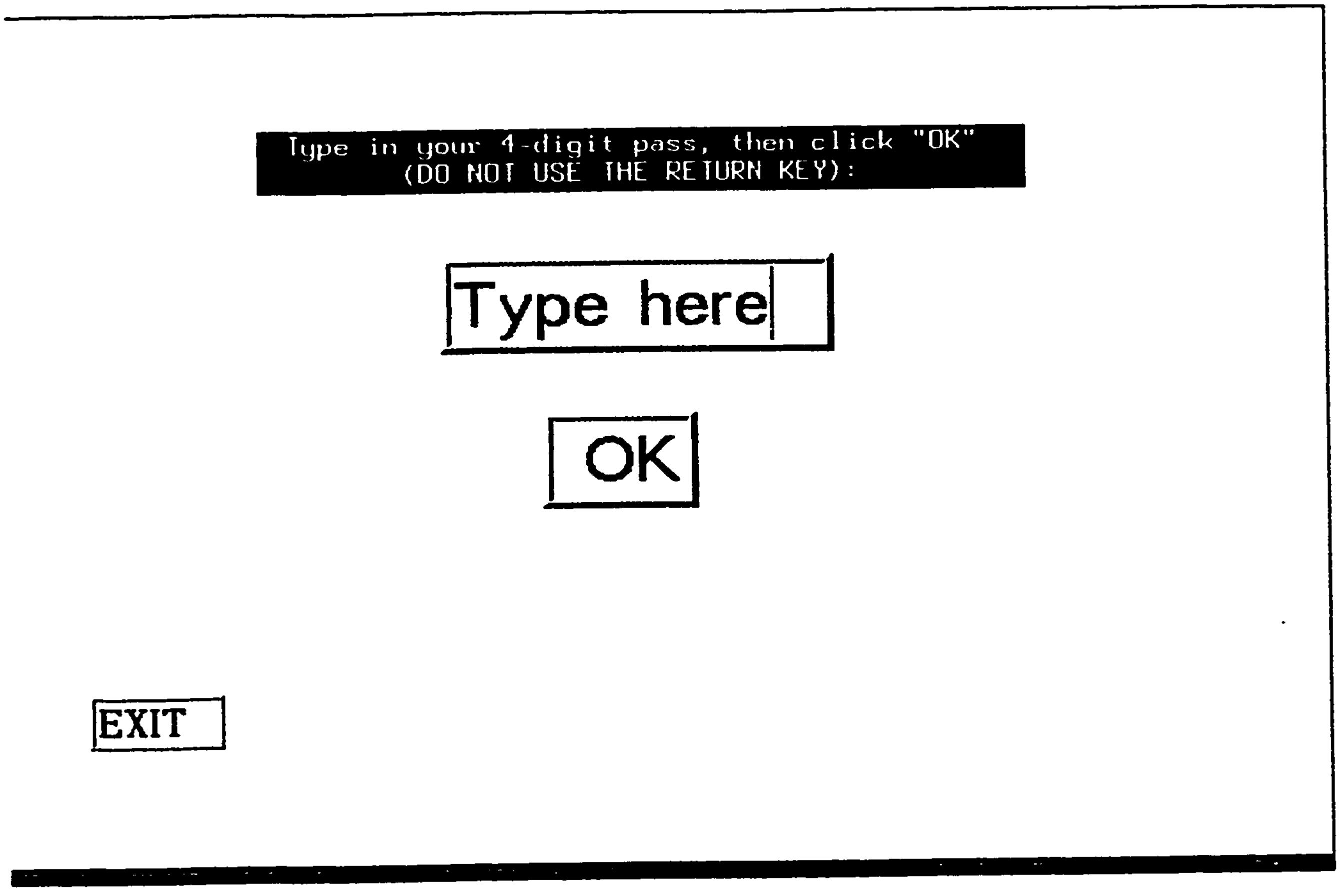
Printout of the Hypertext Module

A STRUCTURAL MODEL OF IMPLEMENTATION

module developed by
A. RAMIREZ

based on
INFORMATION SYSTEMS IMPLEMENTATION:
Testing a Structural Model,
by
Henry C. Lucas, Jr.
Michael J. Ginzberg &
Randall L. Schultz




Type in your 4-digit pass, then click "OK"
(DO NOT USE THE RETURN KEY):
Type here
OK
EXIT

Introduction:

The purpose of this module is to present a rigorous, integrative model of the implementation process.

In particular, the model consists of the adoption of a system by a manager depicting the process by which systems are implemented; managers authorize and approve development and users work with the resulting system. This model, developed and tested by Lucas, Ginzberg and Schultz (90,) includes the major factors and dimensions uncovered in past factor studies of implementation.

Your GOAL in this module is
TO UNDERSTAND THE FACTORS AND DIMENSIONS INCLUDED IN THIS MODEL AS WELL AS TO IDENTIFY THEIR RELATIONSHIPS.

To Module

Information has become an important and sometimes strategic resource in the organization. For this reason, an increasing number of firms are developing computer-based information systems to produce and manage this important resource. The development of information systems, however, is not an easy task. Failure is not uncommon, especially in large and complex systems. Hence, it is important to know how to manage the risk associated with the development of new information systems.

The sucess of information systems implementation can be assessed/predicted by evaluating the following dimensions:

- Manager belief in system concept
- Manager knowledge of system
- Manager assessment of system and support
- Manager involvement
- Manager acceptance

Manager belief in system concept

Manager belief in system concept

This dimension measures the extent to which managers believe in the underlying concept or approach behind a system, that is, their belief in the potential of that approach for solving the organization's information or decision problems. We expect that a stronger belief in this concept, will result in a greater incentive for the manager to become involved in systems development and to learn about the system.

One factor, Top management support exerts a direct influence on Manager belief in system concept.

Dimensions

Manager involvement

Manager belief in system concept

This dimension measures the extent to which managers believe in the underlying concept or approach behind a system, that is, their belief in the potential of that approach for solving the organization's information or decision problems. We expect that

<- CLICK HERE TO CLOSE THIS WINDOW

Top management support

This factor measures the level of support exhibited by top management in the organization for the use of computer-based systems in general as well as for a particular system or system concept. Greater top management support should result in managers being more willing to become involved in system development and having greater belief in the system concept.

Manager involvement

This dimension measures the degree (both quantity and quality) of interaction between the manager and the system designer concerning system development. Higher levels of involvement should lead to greater knowledge of the system and more favorable assessments of the system and support.

Factor Top management support exerts a direct influence on Manager involvement.

Manager belief in system concept, a dimension, exerts also a direct influence on Manager involvement.

← Manager belief in system concept

→ Manager knowledge of system

Manager involvement

This dimension measures the degree (both quantity and quality) of interaction between the manager and the system designer concerning system development. Higher levels of involvement should lead to greater knowledge of the system and more

<- CLICK HERE TO CLOSE THIS WINDOW

Top management support

This factor measures the level of support exhibited by top management in the organization for the use of computer-based systems in general as well as for a particular system or system concept. Greater top management support should result in managers being more willing to become involved in system development and having greater belief in the system concept.

Manager knowledge of system

This is a measure of how well a manager understands a particular system. We expect that better understanding of a system's design and capabilities leads directly to increased acceptance.

Factor [Manager decision style] exerts a direct influence on Manager knowledge of system. Also, two dimensions, [Manager belief in system concept,] & [Manager involvement,] influence the manager's knowledge of a system.

← Manager involvement

→ Manager assessment of system and support

Manager assessment of system and support

This dimension measures the manager's evaluation of the quality of the system and its supporting mechanisms (e.g., people, hardware, data). Favorable evaluations should result in increased acceptance.

Dimension Manager involvement exerts a direct influence on Manager assessment of system and support:

Manager knowledge of system

Manager acceptance

Manager acceptance

Acceptance is a predisposition to use a system or its outputs. For a manager, it is a predisposition for others to use a system. This dimension is a measure of the extent to which a manager wants a particular system to be implemented, that is, accepted and used by others.

Factors Manager decision style, Manager demographics, Manager job characteristics, and Organizational support . are expected to exert an influence on manager acceptance, as well as the following two dimensions:

Manager knowledge of system &

Manager assessment of system and support

Manager assessment of system and support

END?

Manager acceptance

Acceptance is a predisposition to use a system or its outputs.
For a manager, it is a predisposition for others to use a system.

<- CLICK HERE TO CLOSE THIS WINDOW

Manager decision style

Decision style refers to the predominant approach a person uses to solve the kinds of problems for which the system is intended. One simple distinction is between analytic and heuristic styles. An analytic decision maker uses a more quantitative approach and formal analysis, while a heuristic decision maker relies more on intuition and experience.
Managers with more analytic styles should be predisposed to accept a computer-based system, while those with more intuitive styles will tend to reject it. In a similar fashion, decision style will likely impact a manager's willingness to learn about a system.

Manager acceptance

<-- CLICK HERE TO CLOSE THIS WINDOW

Manager demographics

Age, time with company and in job, educational background, previous jobs, experience with previous innovations, and so forth may all affect an individual's willingness to accept a system.

Manager acceptance

<- CLICK HERE TO CLOSE THIS WINDOW

Manager job characteristics

This factor is a measure of the task responsibilities of the manager. Different managers have different sets of tasks as their job responsibilities. Some tasks are more amenable to computer-based support than are others. The more a manager's job is comprised of such (supportable) tasks, the more likely he or she is to accept the system.

Manager acceptance

<- CLICK HERE TO CLOSE THIS WINDOW

Organizational Support

This factor measures the degree to which organizational arrangements foster and facilitates access to and use of a system. It includes factors such as availability of terminals and lines, support facilities (like information centers or consulting support), maintenance of software, databases and chargeback for usage. The greater the support, the greater the acceptance.

Manager acceptance

☒ <- CLICK HERE TO CLOSE THIS WINDOW

Manager assessment of system and support

This dimension measures the manager's evaluation of the quality of the system and its supporting mechanisms (e.g., people, hardware, data).
Favorable evaluations should result in increased acceptance.

Dimension [Manager involvement] exerts a direct influence on Manager assessment of system and support:

Manager acceptance

<- CLICK HERE TO CLOSE THIS WINDOW

Manager assessment of system and support

<- CLICK HERE TO CLOSE THIS WINDOW

Manager involvement

This dimension measures the degree (both quantity and quality) of interaction between the manager and the system designer concerning system development. Higher levels of involvement should lead to greater knowledge of the system and more favorable assessment of the system and support.

Manager belief in system concept exerts a direct influence on Manager involvement.

Top management support, a dimension, exerts also a direct influence on Manager involvement.

Manager acceptance

<- CLICK HERE TO CLOSE THIS WINDOW

Manager assessment of system and support

<- CLICK HERE TO CLOSE THIS WINDOW

Manager involvement

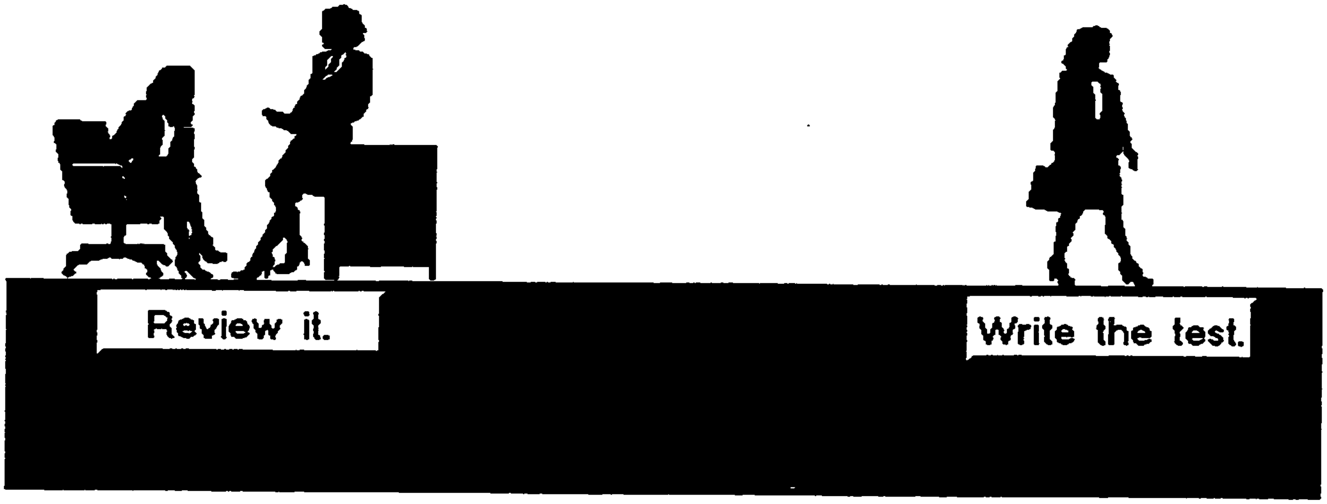
<- CLICK HERE TO CLOSE THIS WINDOW

Top management support

This factor measures the level of support exhibited by top management in the organization for the use of computer-based systems in general as well as for a particular system or system concept. Greater top management support should result in managers being more willing to become involved in system development and having greater belief in the system concept.

Congratulations!

You have finished "**THE MANAGER MODEL.**"



Appendix B

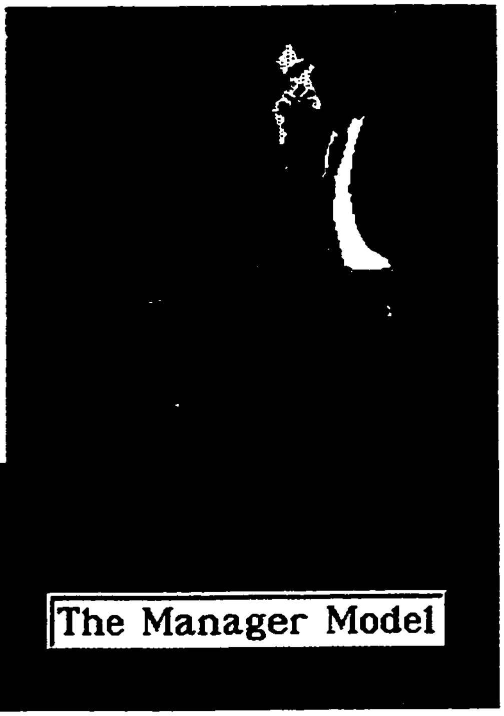
Printout of the Linear text Module

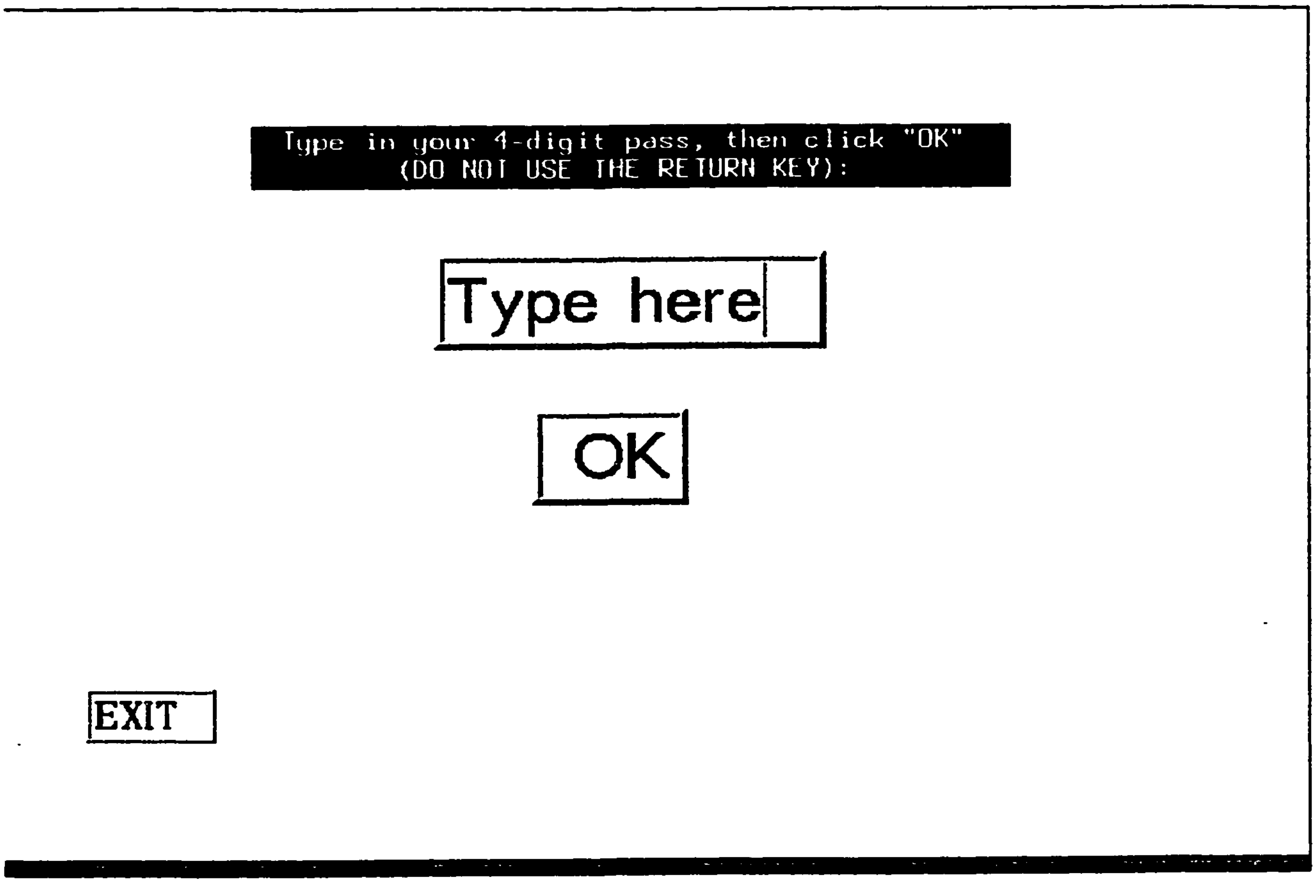
A STRUCTURAL MODEL OF IMPLEMENTATION

module developed by
A. RAMIREZ

based on
INFORMATION SYSTEMS IMPLEMENTATION:
Testing a Structural Model,
by

Henry C. Lucas, Jr.
Michael J. Ginzberg &
Randall L. Schultz




Type in your 4-digit pass, then click "OK"
(DO NOT USE THE RETURN KEY):
Type here
OK
EXIT

Introduction:

The purpose of this module is to present a rigorous, integrative model of the implementation process.

In particular, the model consists of the adoption of a system by a manager depicting the process by which systems are implemented; managers authorize and approve development and users work with the resulting system. This model, developed and tested by Lucas, Ginzberg and Schultz (90,) includes the major factors and dimensions uncovered in past factor studies of implementation.

Your GOAL in this module is
TO UNDERSTAND THE FACTORS AND DIMENSIONS INCLUDED IN THIS MODEL AS WELL AS TO IDENTIFY THEIR RELATIONSHIPS.

To Module

Information has become an important and sometimes strategic resource in the organization. For this reason, an increasing number of firms are developing computer-based information systems to produce and manage this important resource. The development of information systems, however, is not an easy task. Failure is not uncommon, especially in large and complex systems. Hence, it is important to know how to manage the risk associated with the development of new information systems.

The sucess of information systems implementation can be assessed/predicted by evaluating the following dimensions: Manager belief in system concept, Manager knowledge of system, Manager assessment of system and support, Manager involvement, and Manager acceptance.

Manager belief in system concept

Manager belief in system concept

This dimension measures the extent to which managers believe in the underlying concept or approach behind a system, that is, their belief in the potential of that approach for solving the organization's information or decision problems. We expect that a stronger belief in this concept, will result in a greater incentive for the manager to become involved in systems development and to learn about the system.

One factor, Top management support, exerts a direct influence on Manager belief in system concept.

Dimensions

Top management support

Top management support

This factor measures the level of support exhibited by top management in the organization for the use of computer-based systems in general as well as for a particular system or system concept. Greater top management support should result in managers being more willing to become involved in system development and having greater belief in the system concept.

← Manager belief in system concept

→ Manager involvement

Manager involvement

This dimension measures the degree (both quantity and quality) of interaction between the manager and the system designer concerning system development. Higher levels of involvement should lead to greater knowledge of the system and more favorable assessments of the system and support.

Factor: Top management support exerts a direct influence on Manager involvement.

Manager belief in system concept, a dimension, exerts also a direct influence on Manager involvement.

Top management support

Manager decision style

Manager decision style

Decision style refers to the predominant approach a person uses to solve the kinds of problems for which the system is intended. One simple distinction is between analytic and heuristic styles. An analytic decision maker uses a more quantitative approach and formal analysis, while a heuristic decision maker relies more on intuition and experience.
Managers with more analytic styles should be predisposed to accept a computer-based system, while those with more intuitive styles will tend to reject it. In a similar fashion, decision style will likely impact a manager's willingness to learn about a system.

← Manager involvement

→ Manager knowledge of system

Manager knowledge of system

This is a measure of how well a manager understands a particular system. We expect that better understanding of a system's design and capabilities leads directly to increased acceptance.

Factor Manager decision style exerts a direct influence on Manager knowledge of system.
Also, two dimensions: Manager belief in system concept, and Manager involvement, influence the manager's knowledge of a system.

← Manager decision style

→ Manager assessment of system and support

Manager assessment of system and support

This dimension measures the manager's evaluation of the quality of the system and its supporting mechanisms (e.g., people, hardware, data). Favorable evaluations should result in increased acceptance.

Dimension Manager involvement exerts a direct influence on Manager assessment of system and support.

← Manager knowledge of system

→ Manager acceptance

Manager acceptance

Acceptance is a predisposition to use a system or its outputs. For a manager, it is a predisposition for others to use a system. This dimension is a measure of the extent to which a manager wants a particular system to be implemented, that is, accepted and used by others.

Factors: Manager decision style, Manager demographics, Manager job characteristics, and Organizational Support are expected to exert an influence on manager acceptance, as well as the following two dimensions:

Manager knowledge of system and

Manager assessment of system and support.

Manager assessment of system and support

Manager job characteristics

Manager job characteristics

This factor is a measure of the task responsibilities of the manager. Different managers have different sets of tasks as their job responsibilities. Some tasks are more amenable to computer-based support than are others. The more a manager's job is comprised of such (supportable) tasks, the more likely he or she is to accept the system.

← Manager acceptance

→ Manager demographics

Manager demographics

Age, time with company and in job, educational background, previous jobs, experience with previous innovations, and so forth may all affect an individual's willingness to accept a system.

Manager job characteristics

Organizational Support

Organizational Support

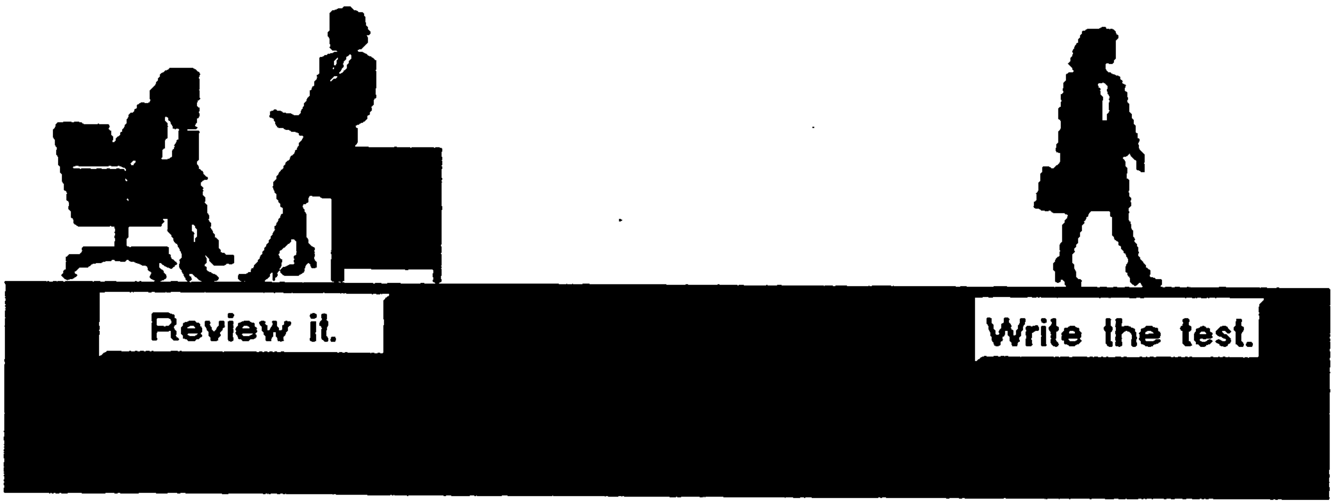
This factor measures the degree to which organizational arrangements foster and facilitates access to and use of a system. It includes factors such as availability of terminals and lines, support facilities (like information centers or consulting support), maintenance of software, databases and chargeback for usage. The greater the support, the greater the acceptance.

Manager demographics

End!

Congratulations!

You have finished "**THE MANAGER MODEL.**"



Appendix C

Printout of the Training Module

TRAINING MODULE

Click here to begin

TRAINING MODULE

In this module you will review some characteristics of the learning module:

* ICONS are defined and you will be able to use them.

* How to close several windows at the same time.

* How to close them one at a time.

* The arrows at the bottom are graphical links.

* Buttons & Highlighted words are also links.

TRY THEM ALL!!!

CLICK ANYWHERE IN THIS SCREEN TO CLOSE IT

The little pictures at the top of the page are called ICONS.

Home Page!

Whenever you want to come back
to this page, click on the X ICON

You will be able to jump back here immediately, without
having to move around the other pages in the module!

TRY IT!

Click on TRY ICON and come back
to this page by clicking on the X ICON.

Click on here to go to point A

TRY ICON

Click on here to go to point B

POINT C

The little pictures at the top of the page are called ICONS.

Home Page!

Whenever you want to come back to this page, click on the X ICON

You will be able to jump back here immediately, without having to move around the other pages in the module!

Click on here to go to point A

Click on here to go to point B

POINT A

POINT A

<- Click here to close this window

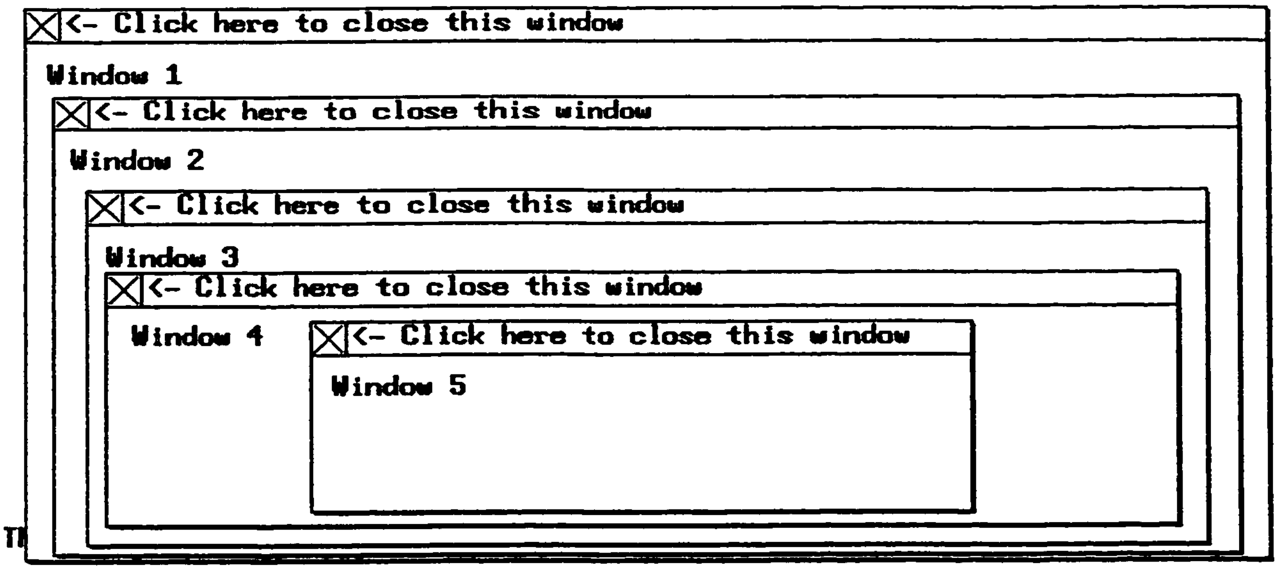
To close one window,
click on the X at the Top Left Corner.

Let's do it!!!

POINT A

If you want to close several windows at the same time use the icon.

Try it now to close all these windows!



POINT B:
Here you will have the opportunity to learn about colour codes to indicate whether or not you have visited a specific location.

1 Click here to open this window

2 Click here to open this window

3 Click here to open this window

THIS ARROW TAKES YOU TO POINT A

THIS ARROW TAKES YOU TO THE END!

POINT B:
Here you will have the opportunity to learn about colour codes to indicate whether or not you have visited a specific location.

1 Click here to open this window

2 C

3 C

<- Click here to close this window

If you want, click on the
ICON to close them.

THIS ARROW TAKES YOU TO POINT A

THIS ARROW TAKES YOU TO THE END!

END?

Congratulations!

You have finished **THE TRAINING MODULE.**

Would you like to review it again ?

Yes

No

Appendix D

Printout of Testing Material

Instructions:

The test contains several multiple choice questions. Read each question carefully. Then, choose the option that best describes or contains the information sought.

For the two cases presented later on, you do not have to memorize any of the information. From the questions there is a link to it. If you want, you can review the case after you have read the questions.

Please note, you will not be able to review a question. Once you have chosen an option, you will work on the next one.

I'd like to review the module one more time

I'm ready to write the test

Question 1: [1 point]

Which one of the following variables is NOT a DIMENSION in the Model?

- Manager acceptance
- Manager assessment of system & support
- Manager belief in system concept
- Manager decision style

Question 2: [1 point]

Which one of the following variables is a DIMENSION in the Model?

Manager decision style

Manager demographics

Manager involvement

Manager job characteristics

Question 3: [1 Point]

Which of the following dimensions is the central variable in the manager model; i.e.; the predisposition to use a system or its outputs, and a measure of the extent to which a manager wants a particular system to be implemented.

- Manager knowledge of system
- Manager involvement
- Manager assessment of system and support
- Manager acceptance
- Manager belief in system concept

Question 4: [2 Points]

Only one of the following variables does NOT have a DIRECT influence on MANAGER ACCEPTANCE. Which is it?

- Manager assessment of system & support
- Manager knowledge of system
- Manager involvement
- Manager job characteristics

Question 5: [2 Points]

Only one of the following variables has a DIRECT influence on MANAGER ACCEPTANCE. Which is it?

- Manager decision style
- Top management support
- Manager involvement
- Manager belief in system concept

Question 6: [3 Points]

The definition of MANAGER BELIEF IN SYSTEM CONCEPT is:
"This variable measures the extent to which a manager believes in the underlying concept or approach behind a system, that is, his or her belief in the potential of that approach for solving the organization's information or decision problems.
We expect that stronger belief in the system concept will result in greater incentive for the manager to become involved in system development and to learn about the system."

From this definition we can infer that MANAGER BELIEF IN SYSTEM CONCEPT will have a DIRECT influence on:

Manager knowledge of system

Manager assessment of system & support

Top management support

Manager acceptance

Integration of Information Technology and Corporate Strategy at Lithonia

Lithonia is the largest American manufacturer of commercial lighting products, with 1989 sales of $659 million. Part of this success is due to Light Link, a networked computer system that puts a terminal, personal computer, software, and the ability to dial into Lithonia at each user's fingertips. Lithonia's seven dozen agents are independent operators who act as intermediaries between the manufacturer and most customers. Light Link connects its warehouses and field sales teams with the agents, specifiers, contractors, and other people who buy, resell or install lighting products. It took seven years and $20 million to develop Light Link, but when it was developed in 1986, the response was immediate. Using the system, independent lighting supply agents could now dial directly into Lithonia's computers for information on inventory, delivery dates, pricing, and other matters. The amount of time that agents spent on the phone with Lithonia representatives plunged, leaving them more time to sell.

This solution was possible because the company's senior vice-president of MIS and general manager of three of its product divisions understood all aspects of the lighting business.

Integration of Information Technology and Corporate Strategy at Lithonia

Question 7: [3 Points]

One conclusion we can infer from the previous information is that the system was accepted because the manager...

- has a positive assessment of the system and its support.
- believes in the system concept.
- got involved in the development of the system.
- knows the system.

Integration of Information Technology and Corporate Strategy at Lithonia

Question 8: [3 Points]

One conclusion we cannot infer from the previous information is about

- the organizational support for Light Link.
- the Manager's decision style.
- Top management support for Light Link.
- the Manager's job characteristics.

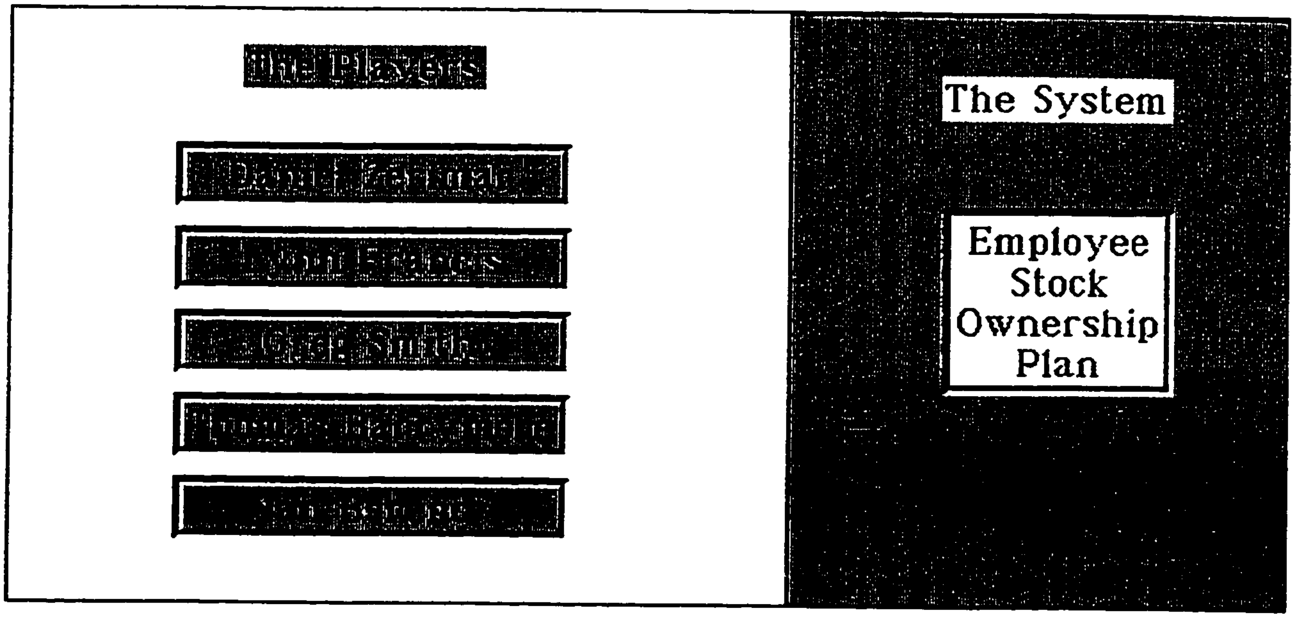
CASE: [20 Points]
The case information is a group of facts organized in six parts: five players and one system. Before proceeding to the questions please read the information in each one of the parts!

You do not have to memorize any of the information.
From the questions there is a link to it. If you want, you can review the case after you have read the questions.

What follows are some facts related to ESOP [Employee Stock Ownership Plan], an Information System, developed and implemented in a particular organization, which

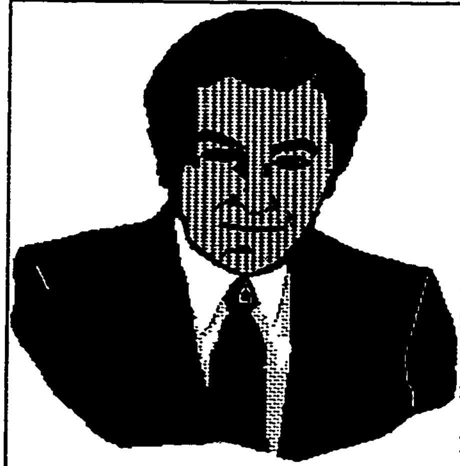
* Is a major services company with offices throughout the United States and Europe.
* Places a great deal of emphasis on human assets.
* Employees receive many benefits from this firm.
* Its information systems department is extremely user-oriented.

AN EVOLUTIONARY SYSTEM



TEST

Daniel Zerimar



* MBA. [Harvard Graduate.]
* Treasurer.
* Main user of the system.
* Joined the company after seven successful years with another company. [He was invited to join mainly due to his experience with similar systems.]

* Has clear goals and a style that invites participation and cooperation
* Liked ESOP and, immediately after joining the company, had several meetings with Lynn Francis and Greg Smith.

* Main objective during these meetings was to fully understand ESOP's characteristics.
* Has mentioned on several occasions that the computer is ESOP's natural "habitat."

TEST

Daniel Zerimar

Lynn Francis

* Implementor
* Developed and updates the file maintenance program containing the information on employee benefits. It has been expanded beyond the twenty individuals originally included.

* Held numerous meetings with Greg Smith and the committee of the top financial staff of the firm to develop the objectives for the "key man" system and the assumptions and characteristics of the projection model.

* Reviewed the calculations with this committee because its members were often surprised by the results from test runs.

CLOSE THIS WINDOW

TEST

Daniel Zerimar

Lynn Francis

Greg Smith

* Manager of the Information Services Department.
* Wrote a small program to perform calculations based on projection rules and assumptions provided by Thomas Barrymore.
* Validated the model using a calculator.
* Developed a rough design for two file records and several formats for screens on a monitor.

* Gave it to Lynn Francis.

* Loaded stock ownership records for all employees into the database.

* Refers to his personnel as "implementors."

CLOSE THIS WINDOW

Daniel Zerimar

Lynn Francis

Greg Smith

Thomas Barrymore

* Vice Chair of the Board
* First to suggest ESOP to Ian Berend
* Asked for a small model to predict the impact of an Employee Stock Ownership Plan.

CLOSE THIS WINDOW

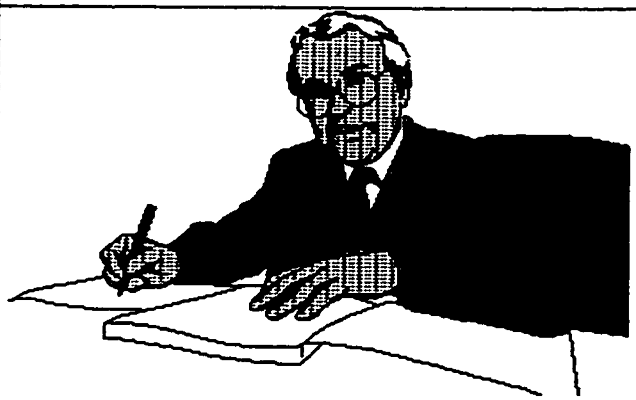
Daniel Zerimar

Lynn Francis

Greg Smith

Thomas Barrymore

Ian Berend

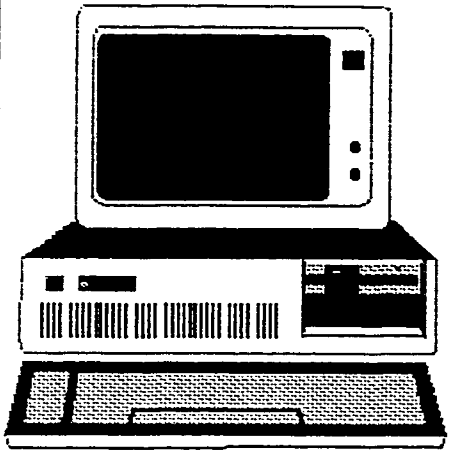


* Manager
* Senior Vice President
* Member of the Finance Committee, (one of the most important committees in the management of the firm.)

CLOSE ALL WINDOWS

TEST

ESOP = Employee Stock Ownership Plan



* The model was designed to answer questions such as: If we adopt an ESOP, how many shares of stock will be needed by the company in ten, twenty and thirty years? What level of growth is necessary to support the ESOP requirements?
* This model raised questions about the impact of the ESOP, which appeared greater than originally anticipated.
* The model was validated using a calculator.

TEST

ESOP = Employee Stock Ownership Plan

* The model did accurately reflect the rules and assumptions supplied.
* The output of this model was presented to the executive committee [ESOP was adopted partially based on this information.]
* ESOP's conversion was a gradual process as various parts of the system were completed.
* Currently staff members in the treasury and accounting area enter and maintain employee records on stock ownership.
* Occasional requests occur from different executives for projections on individuals.
* The system expanded to include all the employees' stock records.
* The system is gradually being extended to include all employee benefits' data.
* A major change was in top management perceptions on compensation equity.

TEST

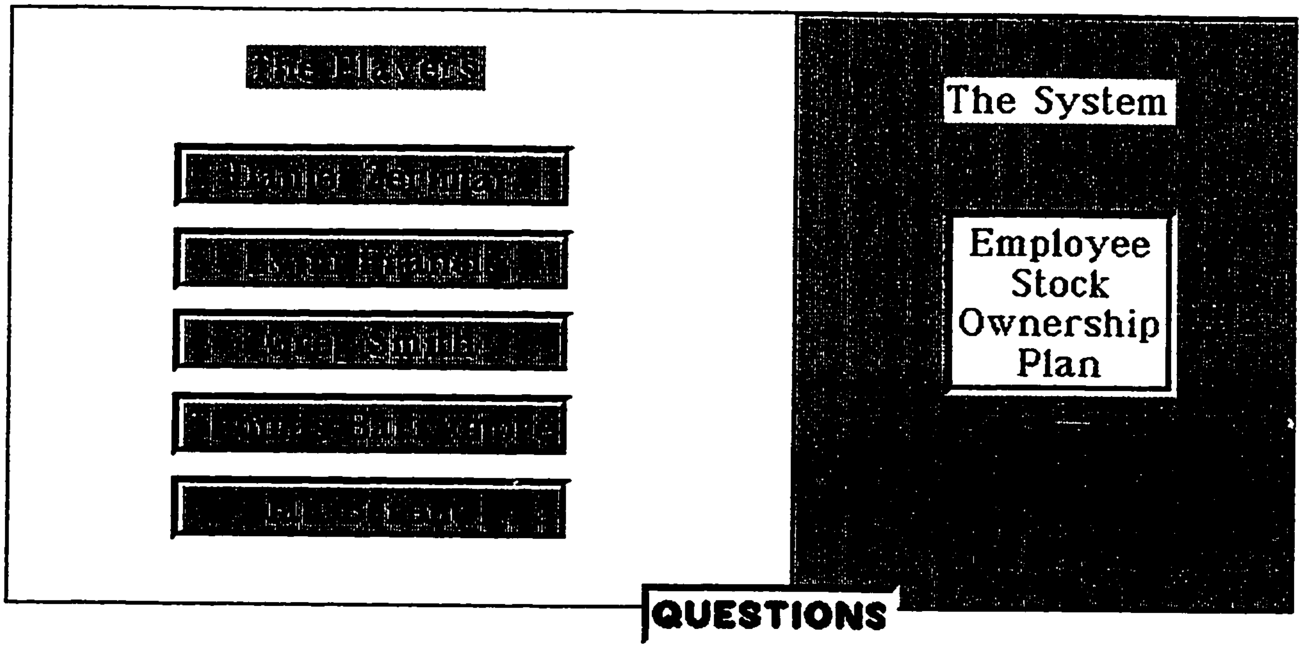
ESOP = Employee Stock Ownership Plan

* The model did accurately reflect the rules and assumptions supplied.
* The output of this model was presented to the executive committee [ESOP was adopted partially based on this information.]
* ESOP's conversion was a gradual process as various parts of the system were completed.
* Currently staff members in the treasury and accounting area

* The information system helped the development of a more equitable compensation program.
* The system also stimulated the creation of a new corporate compensation committee.
* The system helped the accomplishment of a top management goal: moving forward more centralized financial processing for better quality control and efficiency.

TEST

AN EVOLUTIONARY SYSTEM



Question 9: [5 Points]

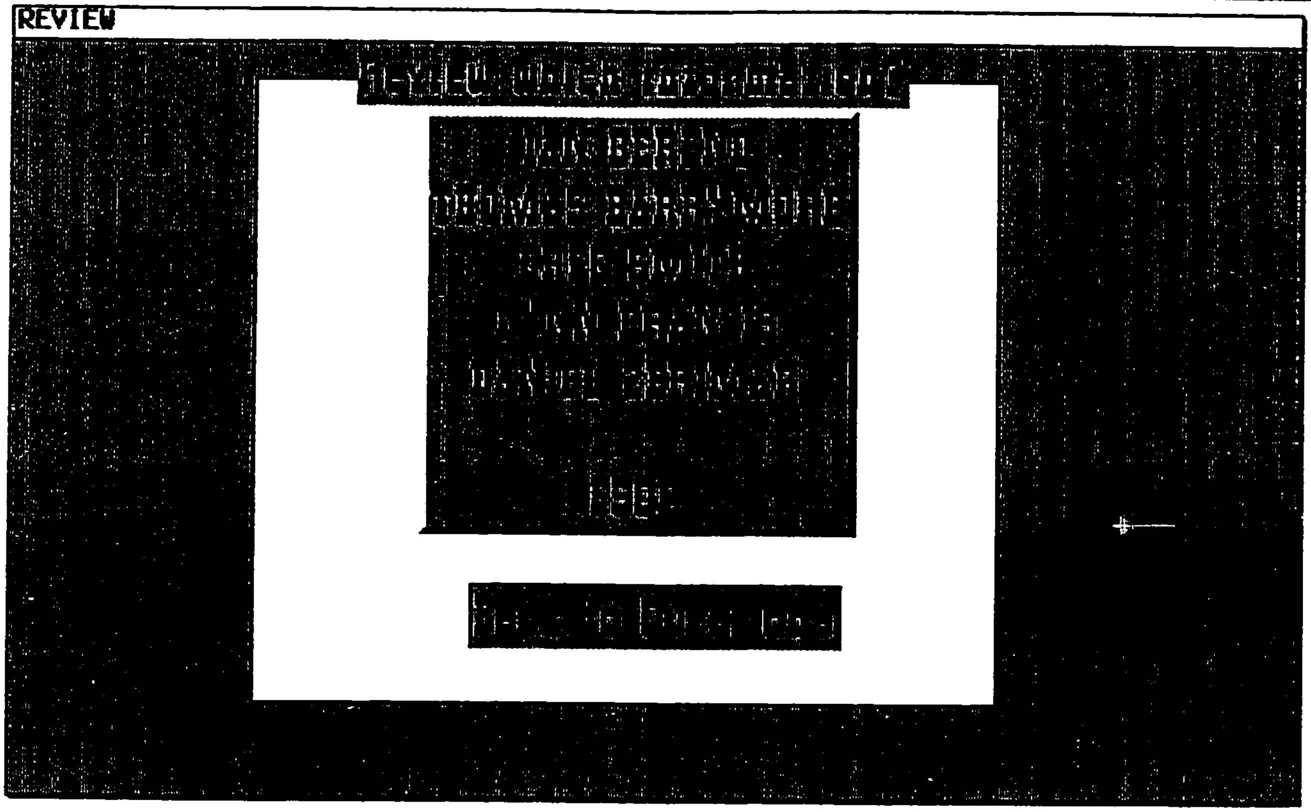
ESOP was successfully implemented mainly due to:

Mr. Zerimar's style. Being a Harvard graduate he liked it and decided to implement it regardless of its cost.

Mr. Berend's support for Mr. Barrymore's idea fostered an organizational need for it.

Mr. Zerimar's acceptance. He thinks that the organization was ready for it. He believes in those systems and knows ESOP quite well due to his involvement.

Mr. Zerimar's need for a tool to show the impact of ESOP to all the members of the finance committee.


REVIEW

Question 10: [3 Points]

Do you think that Mr. Zerimar believes in the system concept?

Yes. Mr. Zerimar believes that these kind of systems are better to manage information such as ESOP's.

No. Mr. Zerimar has a positive assessment of ESOP mostly due to Mr. Berends support for it.

Yes. Mr. Zerimar believes in any system. He learnt to like these systems in Harvards School of Management.

No. not really. what Mr. Zerimar aknowledges is the support that the organization has for ESOP.

Question 11: [4 Points]

Which of the following statements describes better Mr. Zerimar's assessment of ESOP and its support?

Mr. Zerimar's knows that the system has great support but he's not quite satisfied with ESOP's quality.

Mr. Zerimar likes ESOP even though he knows that he'll have to work hard to have the organizational support he wants.

Mr. Zerimar joined the company mainly due to ESOP. He knew that the organization was ready for such a system.

There is not enough information in the case to determine Mr. Zerimar's assessment of ESOP and its support.

Question 12: [4 Points]

Do you think that Mr. Zerimar's acceptance of ESOP is based more on his knowledge of it than in his involvement in its development?

Neither. Mr. Zerimar's acceptance is based solely on his preference for these kinds of systems.

No, it is the other way around. Mr. Zerimar's acceptance is based on his involvement in ESOP's development.

Yes, his knowledge of ESOP is the main reason he accepted it.

Yes, his knowledge of ESOP has a direct influence on his acceptance while his involvement has only an indirect influence.

Question 13: [4 Points]

Do you think that Mr. Zerimar's MBA, position and experience directly influenced his acceptance of ESOP?

- His position and experience yes, but not his degree.
- No, his acceptance was independent of any of these factors.
- Yes, all these factors directly influenced his acceptance.
- No, these factors affect some dimensions but not acceptance.

THE TEST HAS ENDED! WHAT FOLLOWS IS AN EVALUATION
OF YOUR EXPERIENCE IN THIS EXPERIMENT,
TO CONTINUE CLICK ON EVALUATION.

EVALUATION

Thank you for your participation in this Project.

Appendix E

Printout of the Evaluation Form

EVALUATION FORM

Thank you for your participation in this research.
Without you, it would have not been possible.
As a member of one of the two groups,
your experience is unique.
We would like to know more about it.

Please take the time to give us your opinion.

1.- About my participation in this research:

I am glad I participated	1	2	3	4	5	6	7	I am not glad I participated

CHOOSE THE NUMBER THAT BEST DESCRIBES HOW YOU FEEL ABOUT THIS EXPERIENCE

2.- About the learning material (content):

I liked it	1	2	3	4	5	6	7	I didn't like it.

CHOOSE THE NUMBER THAT BEST DESCRIBES HOW YOU FEEL ABOUT THIS EXPERIENCE

2.- About the learning material (content):

I found it useful	1	2	3	4	5	6	7	I did not find it useful

CHOOSE THE NUMBER THAT BEST DESCRIBES HOW YOU FEEL ABOUT THIS EXPERIENCE

2.- About the learning material (content):

I knew it | 1 | 2 | 3 | 4 | 5 | 6 | 7 | I didn't know it

CHOOSE THE NUMBER THAT BEST DESCRIBES HOW YOU FEEL ABOUT THIS EXPERIENCE

3.- About the method I was assigned to work:
[Either Linear or Hypertext]

I liked it	1	2	3	4	5	6	7	I didn't like it.

CHOOSE THE NUMBER THAT BEST DESCRIBES HOW YOU FEEL ABOUT THIS EXPERIENCE

3.- About the method I was assigned to work:
[Either Linear or Hypertext]

I enjoyed it	1	2	3	4	5	6	7	I didn't enjoy it

CHOOSE THE NUMBER THAT BEST DESCRIBES HOW YOU FEEL ABOUT THIS EXPERIENCE

3.- About the method I was assigned to work:
[Either Linear or Hypertext]

It was friendly	1	2	3	4	5	6	7	It was unfriendly

CHOOSE THE NUMBER THAT BEST DESCRIBES HOW YOU FEEL ABOUT THIS EXPERIENCE

4.- About the test:

It was too easy	1	2	3	4	5	6	7	It was too difficult

CHOOSE THE NUMBER THAT BEST DESCRIBES HOW YOU FEEL ABOUT THIS EXPERIENCE

4.- About the test:

Cases were interesting	1	2	3	4	5	6	7	Cases were not interesting

CHOOSE THE NUMBER THAT BEST DESCRIBES HOW YOU FEEL ABOUT THIS EXPERIENCE

4.- About the test:

Cases were challenging	1	2	3	4	5	6	7	Cases were not challenging

CHOOSE THE NUMBER THAT BEST DESCRIBES HOW YOU FEEL ABOUT THIS EXPERIENCE

5.- About the Hardware:
[Mainly the Computer Mouse]

I have used it before	1	2	3	4	5	6	7	I have never used it before

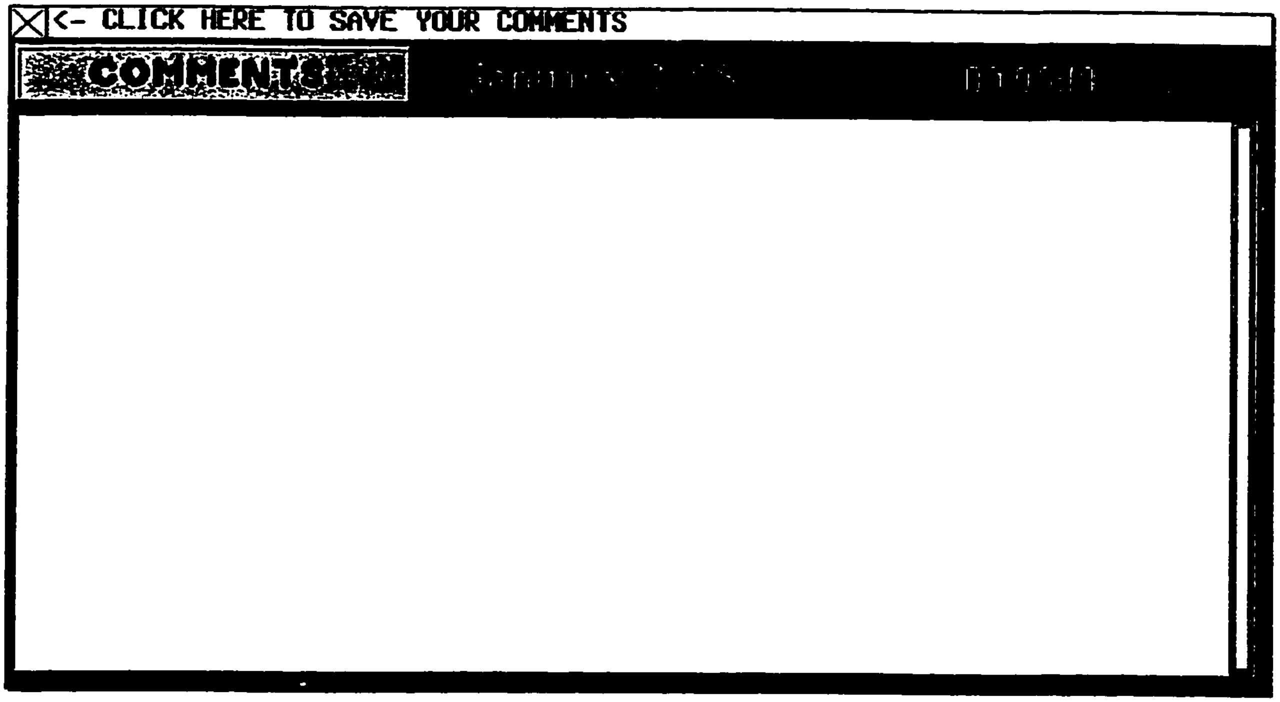
CHOOSE THE NUMBER THAT BEST DESCRIBES YOUR EXPERIENCE
WITH THIS HARDWARE

6.- In general:

I had a wonderful experience	1	2	3	4	5	6	7	I had a terrible experience

CHOOSE THE NUMBER THAT BEST DESCRIBES HOW YOU FEEL ABOUT THIS EXPERIENCE

Please feel free to use this page to add any comment you think will help us improve this research.



Please feel free to use this page to add any comment you think will help us improve this research.

THANK YOU!!!!!!!!!!!!!!!!!!!!!

Now, you can continue

→

Appendix F

Additional Experimental Material

CONCORDIA UNIVERSITY

Faculty of Commerce and Administration

Department of Decision Sciences and Management Information Systems

Thank you for your participation on this research. You will have the opportunity to win **$100.00** and learn a specific Systems Implementation model. This research is testing a new computer-based training and teaching methodology.

You will be randomly assigned to one of the versions. The learning process for each group takes approximately the same time (**60-90 minutes**). The time spent is not part of the research, which means you can do it at your own pace.

There is an introduction to the system for both groups. The introduction gives the opportunity to know the system's characteristics and utilities (basically, how to move around). Once participants are familiar with the system's characteristics, they will be logged into the system containing the learning material. When they consider to be ready, they will continue with the evaluation material.

There is a **$100.00 (One hundred dollars)** ***FIRST PRIZE*** for *each group*. Prizes will be awarded on performance. Performance will be measured with the evaluation material. How to break ties will be decided by those with tied scores. Each one of you will receive a written evaluation of your performance that may help you further on learning similar material.

	RESEARCH SCHEDULE AND LOCATION:
WHEN:	Weekdays from September 29 to October 10, 1997. From 10:00 to 22:00 Hrs. [Entry times: 10:00; 11:30; 14:00; 15:30; 17:30; 19:00 and 20:30 Hrs.]
WHERE:	GM 710-01 Your appointment's time will be sent to you by September 24, 1997 via E-mail.

M.I.S. MODELS

In MIS, as in many other fields, conceptual models are invented to provide an appropriate representation of a system, appropriate in the sense of being accurate, consistent, and complete. These models are invented by researchers, teachers, designers, scientists, and sometimes by MIS practitioners; the goal being to provide predictive and explanatory understanding of the interaction of the different elements of the system.

System Implementation Models are characteristic of the first generation of research on implementation. The ultimate objective in this kind of research is to provide guidelines for the management of implementation. To do so, researchers proceed from an exploratory stage that sets a conceptual foundation, through the definition of variables (usually called dimensions), factors and their relationships, to the integration of these dimensions, factors and relationships into a testable model.

Once the model is tested, if successful, it can be used practically. Results are published and these models are available to the MIS community. Unfortunately, most of these models are limited and explain only part of the implementation process. Since the interest in the problems of implementing information systems keeps on growing, there will be many new models dealing with its dimensions, factors, and their relationships. One thing everybody agrees on is that implementation is the ***raison-d'être*** *of any information systems analysis and design.*

CONCORDIA UNIVERSITY
FACULTY OF COMMERCE AND ADMINISTRATION
Department of Decision Sciences and Management Information Systems

SIGN UP (CONSENT) FORM

NAME: ______________________________ I.D. __________
[PLEASE PRINT] **E-mail:**

PLEASE CHOOSE AT LEAST TWO TIME SLOTS (YOU CAN MARK THEM FIRST AND SECOND CHOICE). FINAL SCHEDULE WILL BE SENT TO YOU BY E-MAIL BY SEPTEMBER 24, 1997. THANK YOU FOR YOUR COOPERATION.

Time Slot	**Monday**	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
10:00 - 11:30					
11:30 - 13:00					
14:00 - 15:30					
15:30 - 17:00					
17:30 - 19:00					
19:00 - 20:30					
20:30 - 22:00					

****_NOTES & COMMENTS_****

PLEASE DO NOT WRITE BELOW THIS LINE

- O **S:A**
- O **M:H**
- O **PA**
- O **G:M**
- O **E-MAILED**
- O **COMPUTERIZED**

CONCORDIA UNIVERSITY
Faculty of Commerce and Administration
Department of Decision Sciences and Management Information Systems

Thank you for your participation on this research. You had the opportunity to win **$100.00** and learn a specific Systems Implementation model. This research tested a new computer-based training and teaching methodology.

You were randomly assigned to one of the two versions. The learning process for each group took approximately the same time: **Mean time in Hypertext was 21 minutes and 37 seconds; Mean time in Linear text was 19 minutes and 59 seconds.**

The top performers in each group were:

Hypertext	First place	M. Misquita	83.3333	Section 51	$100.00
	Second	K. Fiore	77.7778	Section A	$50.00
Linear	First place	W. Chan	83.3333	Section A	$100.00
	Second	A. Lee	77.7778	Section 51	$50.00
	Second	S. Nguyen	77.7778	Section A	$50.00
	Second	N. Duong	77.7778	Section A	$50.00

The Mean of this sample was 52.5556. When we break this information by Method, we get a mean score in Hypertext of 56.0000; a mean score in Linear text of 49.0833. If we break this information by Gender, we get a mean score by Females of 57.9167, a mean score by Males of 48.0556. In this study participated 103 subjects, 52 in Hypertext and 51 in Linear text), 47 Females and 56 Males.

The test measured three levels of complexity: Analysis, Synthesis and Evaluation. Evaluation was measured with a long case study. Synthesis was measured with a short case study. Analysis was measured with a multiple choice test. It was expected that students working on Hypertext will have a greater performance on Evaluation than those working on the Linear module. The mean score on Evaluation was 52.5000. When we break this information by Method, we get a mean score on Evaluation in Hypertext of 57.9000, a mean score on Evaluation in Linear text of 47.0500. If we break this information by Gender, we get a mean score on Evaluation by Females of 59.9000, a mean score on Evaluation by Males of 46.3500.

About Global Satisfaction, this key was measured with a 7-point Likert scale in the evaluation form at the end of the experiment. From 1 "I HAD A WONDERFUL EXPERIENCE" to 7 "I HAD A TERRIBLE EXPERIENCE" this sample considered that they had an experience of 2.35, which is closer to 1 than it is to 7. Those who had the most wonderful experience were Females with an index of 2.04, then those participating in Hypertext with 2.23; after them, those participating in the Linear module with 2.48 and finally were Males with 2.62, still being closer to 1 than to 7.

[Name$_j$], your personal score was [Score$_j$] which is [above|below] the group mean. Breaking this score into its three factors, your performance on Analysis was [Test$_j$] which is [above|below] the group mean. Your performance on Synthesis was [Case-1$_j$] which is [above|below] the group mean. Your performance on Evaluation was [Case-2$_j$] which [above|below] the group mean.

You spent [Time-in-Module$_j$] on the module, [Time-in-Test$_j$] on the test, [Time-in-Case-1$_j$] on case 1, [Time-in-Case-2$_j$] on case 2, with a total time of [Total-Time$_j$].

Thank you for your participation!

[[If Verbal-Protocols.EQ.1 then PRINT: Since you agreed to share your verbal protocols with me working in the Thinking Aloud group you will receive $20.00 (Twenty Dollars).]

[If Score.EQ.Winner then PRINT: Since you are one of the winners, please tell me (via e-mail) when it is convenient for you to meet me and receive your reward.]

Appendix G

Pilot Study Data

Variable	N	Mean	Std. Dev.	Min.	Max
Learning Time	36	17:58	6:17	8:03	35:49
Effectiveness Analysis	36	6.14	3.10	0	10
Effectiveness Synthesis	36	2.83	2.14	0	6
Effectiveness Evaluation	36	10.28	4.27	3	20
Efficiency Analysis	36	55.53%	27.23%	0%	91.86%
Efficiency Synthesis	36	43.25%	33.24%	0%	97.67%
Efficiency Evaluation	36	49.44%	20.43%	14.24%	95.71%
Participation	35	2.09	1.04	1	5
Content Motivation	35	2.66	1.49	1	6
Content Usefulness	35	2.80	1.28	1	6
Content Experience	35	4.66	1.43	2	7
Task Motivation	35	2.91	1.85	1	7
Task Satisfaction	35	2.83	1.69	1	6
Task User Friendliness	35	2.20	1.53	1	7
Test Difficulty	35	4.46	1.09	2	7
Test Challenge	35	2.89	1.16	1	7
Test Motivation	35	3.23	1.52	1	7
Hardware Use	35	1.54	0.92	1	5
Global Satisfaction	35	2.57	1.24	1	6

Table G1: Pilot Study Statistics

Variable	N	Mean	Std. Dev.	Min.	Max
Learning Time	18	17:38	6:56	8:03	35:49
Effectiveness Analysis	18	6.33	3.05	1	10
Effectiveness Synthesis	18	2.33	2.20	0	6
Effectiveness Evaluation	18	11.56	4.09	3	20
Efficiency Analysis	18	57.55%	26.41%	9.46%	90.81%
Efficiency Synthesis	18	35.99%	34.78%	0%	97.67%
Efficiency Evaluation	18	55.72%	18.94%	14.72%	95.71%
Participation	18	1.94	1.00	1	4
Content Motivation	18	2.33	1.41	1	6
Content Usefulness	18	2.44	1.15	1	4
Content Experience	18	4.72	1.41	3	7
Task Motivation	18	2.28	1.23	1	5
Task Satisfaction	18	2.28	1.41	1	6
Task User Friendliness	18	1.56	0.92	1	4
Test Difficulty	18	4.22	0.94	3	6
Test Challenge	18	2.94	1.00	1	4
Test Motivation	18	3.06	1.30	1	5
Hardware Use	18	1.67	1.03	1	5
Global Satisfaction	18	2.56	1.29	1	6

Table G2: Pilot Study Statistics, Method = Hypertext

Variable	N	Mean	Std. Dev.	Min.	Max
Learning Time	18	18:18	5:45	10:14	32:59
Effectiveness Analysis	18	5.94	3.23	0	10
Effectiveness Synthesis	18	3.33	2.03	0	6
Effectiveness Evaluation	18	9.00	4.17	3	16
Efficiency Analysis	18	53.51%	28.65%	0%	91.86%
Efficiency Synthesis	18	50.51%	30.88%	0%	96.70%
Efficiency Evaluation	18	43.16%	20.42%	14.12%	78.63%
Participation	17	2.24	1.09	1	5
Content Motivation	17	3.00	1.54	1	6
Content Usefulness	17	3.18	1.33	1	6
Content Experience	17	4.59	1.50	2	7
Task Motivation	17	3.59	2.18	1	7
Task Satisfaction	17	3.41	1.80	1	6
Task User Friendliness	17	2.88	1.76	1	7
Test Difficulty	17	4.71	1.21	2	7
Test Challenge	17	2.82	1.33	1	7
Test Motivation	17	3.41	1.73	1	7
Hardware Use	17	1.41	0.80	1	4
Global Satisfaction	17	2.59	1.23	1	6

Table G3: Pilot Study Statistics, Method = Linear Text

Appendix H

Experimental Data Descriptive Statistics

Variable	N	Mean	Std. Dev.	Min.	Max
Learning Time	103	20:48	8:56	5:13	54:24
Time in Analysis	103	7:21	2:50	3:03	21:28
Time in Synthesis	103	5:49	3:16	1:57	23:11
Time in Evaluation	103	16:16	6:29	5:35	37:20
Effectiveness Analysis	103	6.00	2.72	0	10
Effectiveness Synthesis	103	2.42	1.93	0	6
Effectiveness Evaluation	103	10.50	4.59	0	20
Efficiency Analysis	103	54.35%	24.52%	0%	100%
Efficiency Synthesis	103	37.18%	30.27%	0%	100%
Efficiency Evaluation	103	50.17%	21.47%	0%	100%
Participation	102	2.13	1.34	1	7
Content Motivation	102	2.47	1.40	1	7
Content Usefulness	102	2.43	1.38	1	7
Content Experience	102	3.98	1.56	1	7
Task Motivation	102	2.44	1.54	1	7
Task Satisfaction	102	2.35	1.36	1	6
Task User Friendliness	102	1.85	1.21	1	7
Test Difficulty	102	4.08	1.08	2	7
Test Challenge	102	2.77	1.06	1	7
Test Motivation	102	2.91	1.45	1	7
Hardware Use	102	1.28	0.67	1	5
Global Satisfaction	102	2.35	1.11	1	6

Table H1: Descriptive Statistics, overall data

Variable	N	Mean	Std. Dev.	Min.	Max
Learning Time	52	21:37	10:30	5:13	54:24
Effectiveness Analysis	52	6.27	2.84	0	10
Effectiveness Synthesis	52	2.31	1.94	0	6
Effectiveness Evaluation	52	11.58	4.41	3	20
Efficiency Analysis	52	56.93%	25.70%	0%	100%
Efficiency Synthesis	52	35.54%	30.78%	0%	98.91%
Efficiency Evaluation	52	55.03%	19.87%	14.72%	95.80%
Participation	52	1.98	1.29	1	7
Content Motivation	52	2.33	1.35	1	7
Content Usefulness	52	2.10	1.18	1	6
Content Experience	52	3.83	1.49	1	7
Task Motivation	52	2.15	1.32	1	7
Task Satisfaction	52	2.04	1.20	1	6
Task User Friendliness	52	1.50	0.75	1	4
Test Difficulty	52	4.00	0.89	2	6
Test Challenge	52	2.81	1.05	1	5
Test Motivation	52	2.62	1.16	1	5
Hardware Use	52	1.25	0.68	1	5
Global Satisfaction	52	2.23	1.13	1	6

Table H2: Descriptive Statistics, Method = Hypertext

Variable	N	Mean	Std. Dev.	Min.	Max
Learning Time	51	19:59	7:01	7:12	41:26
Effectiveness Analysis	51	5.73	2.59	0	10
Effectiveness Synthesis	51	2.53	1.93	0	6
Effectiveness Evaluation	51	9.41	4.55	0	20
Efficiency Analysis	51	51.72%	23.22%	0%	97.75%
Efficiency Synthesis	51	38.86%	29.94%	0%	100%
Efficiency Evaluation	51	45.23%	22.09%	0%	100%
Participation	50	2.28	1.39	1	7
Content Motivation	50	2.62	1.44	1	7
Content Usefulness	50	2.78	1.50	1	7
Content Experience	50	4.14	1.63	1	7
Task Motivation	50	2.74	1.70	1	7
Task Satisfaction	50	2.68	1.45	1	6
Task User Friendliness	50	2.22	1.46	1	7
Test Difficulty	50	4.16	1.25	2	7
Test Challenge	50	2.74	1.08	1	7
Test Motivation	50	3.22	1.66	1	7
Hardware Use	50	1.32	0.65	1	4
Global Satisfaction	50	2.48	1.09	1	6

Table H3: Descriptive Statistics, Method = Linear text

Appendix I

Pre and Post transformations Plots
for the time variables,
Normal Probability Plots
Detrended Normal Probability Plots
(Post transformations)

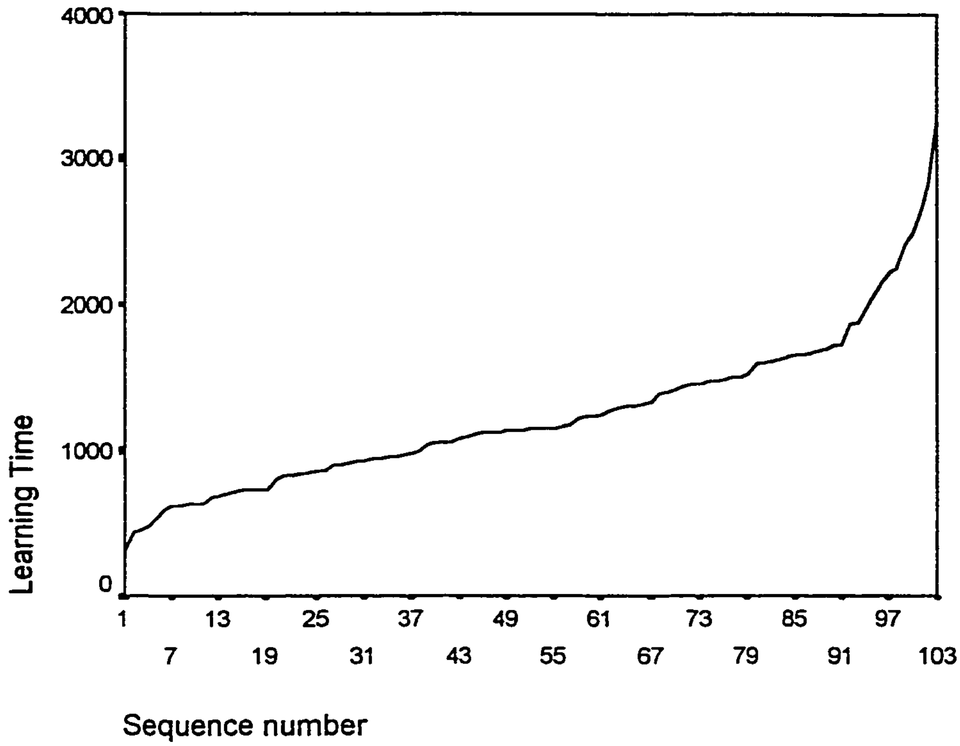


Figure I1: Plot of variable Learning Time [Pre transformation]

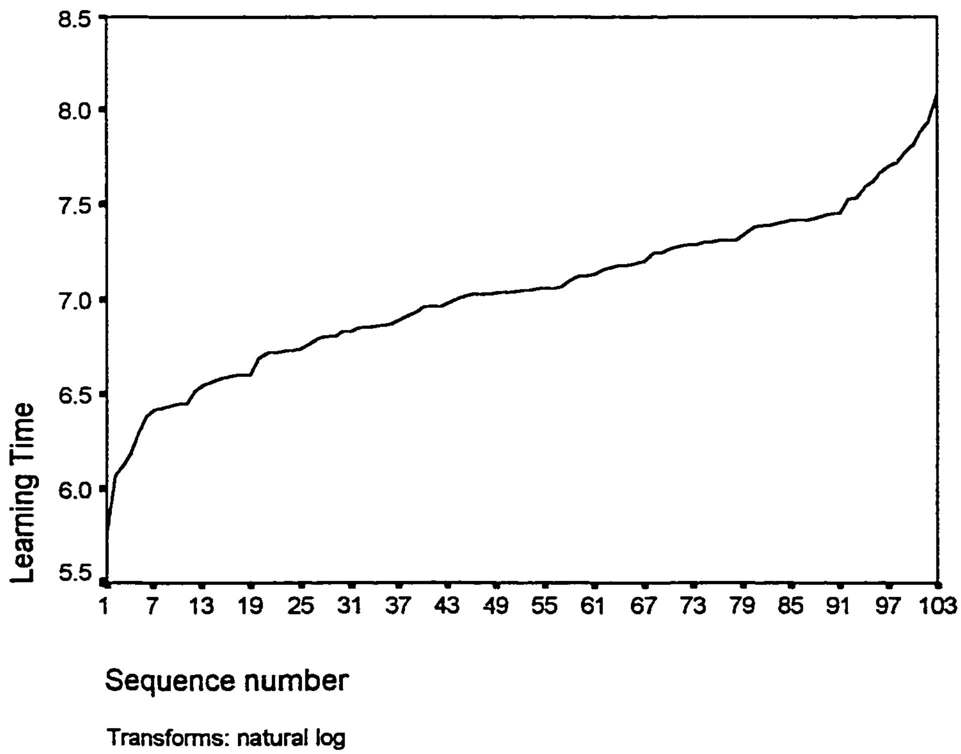


Figure I2: Plot of variable Learning Time [Post transformation]

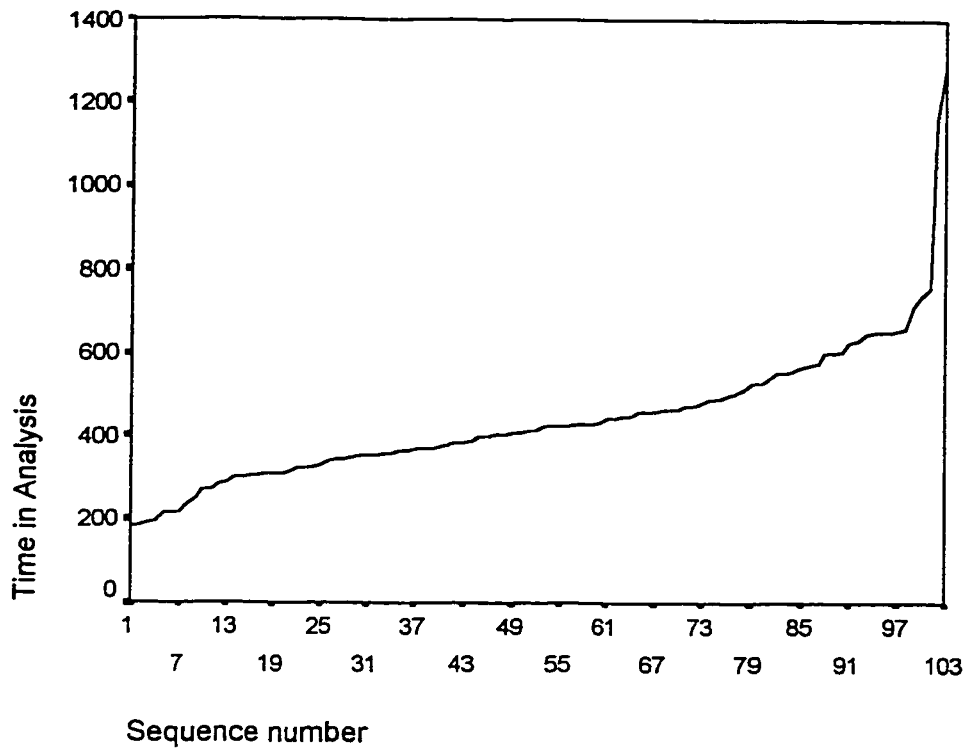


Figure I3: Plot of variable Time in Analysis [Pre transformation]

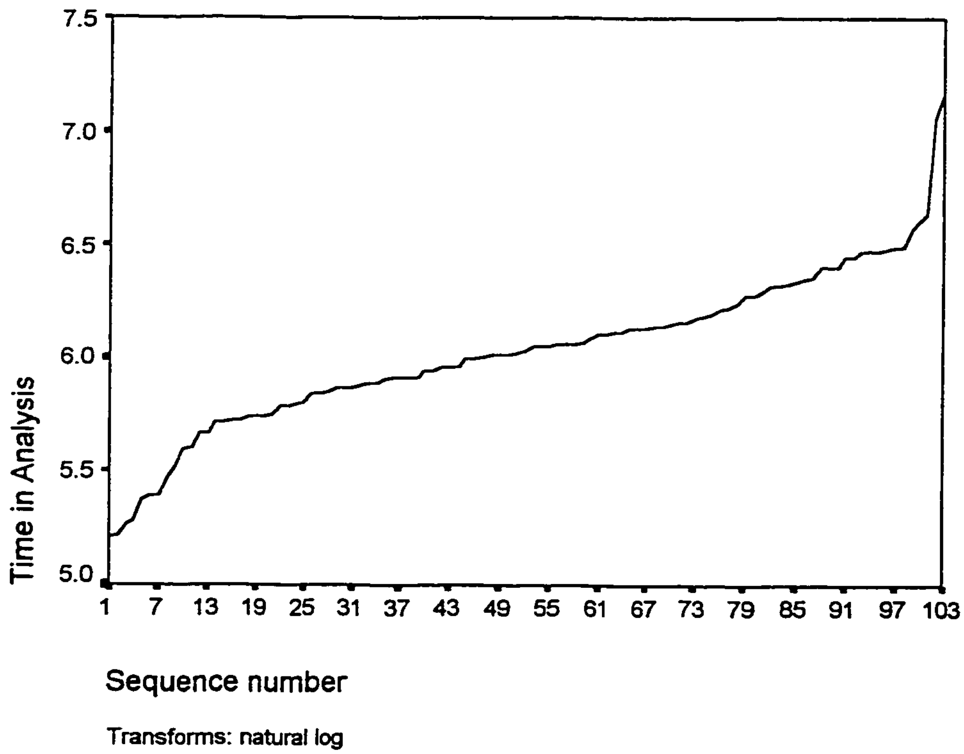


Figure I4: Plot of variable Time in Analysis [Post transformation]

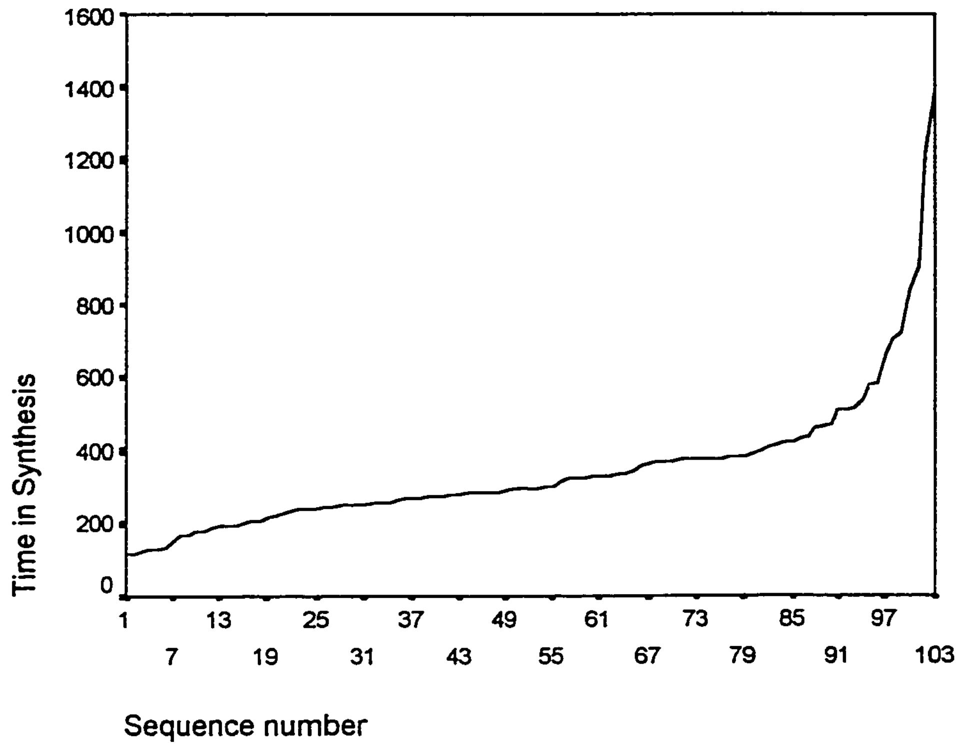


Figure I5: Plot of variable Time in Synthesis [Pre transformation]

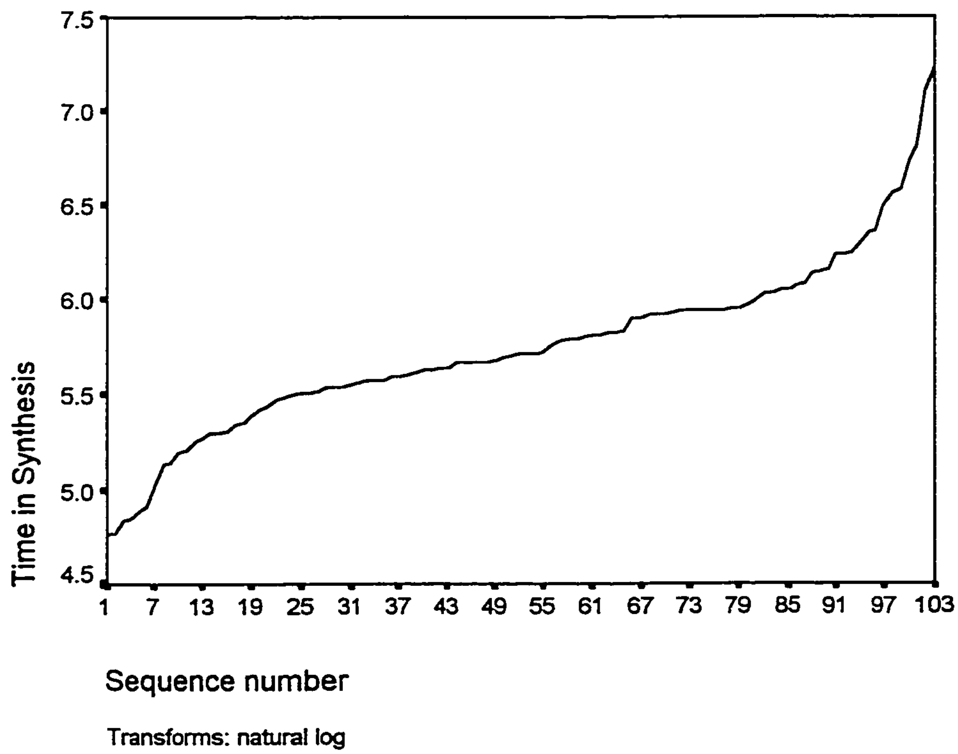


Figure I6: Plot of variable Time in Synthesis [Post transformation]

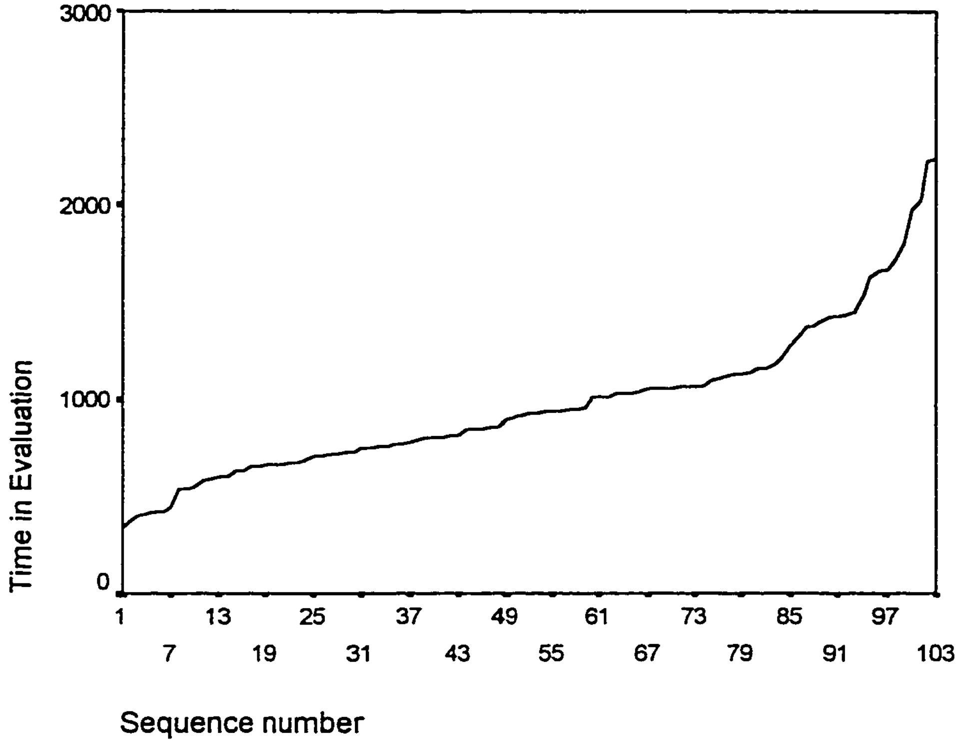


Figure I7: Plot of variable Time in Evaluation [Pre transformation]

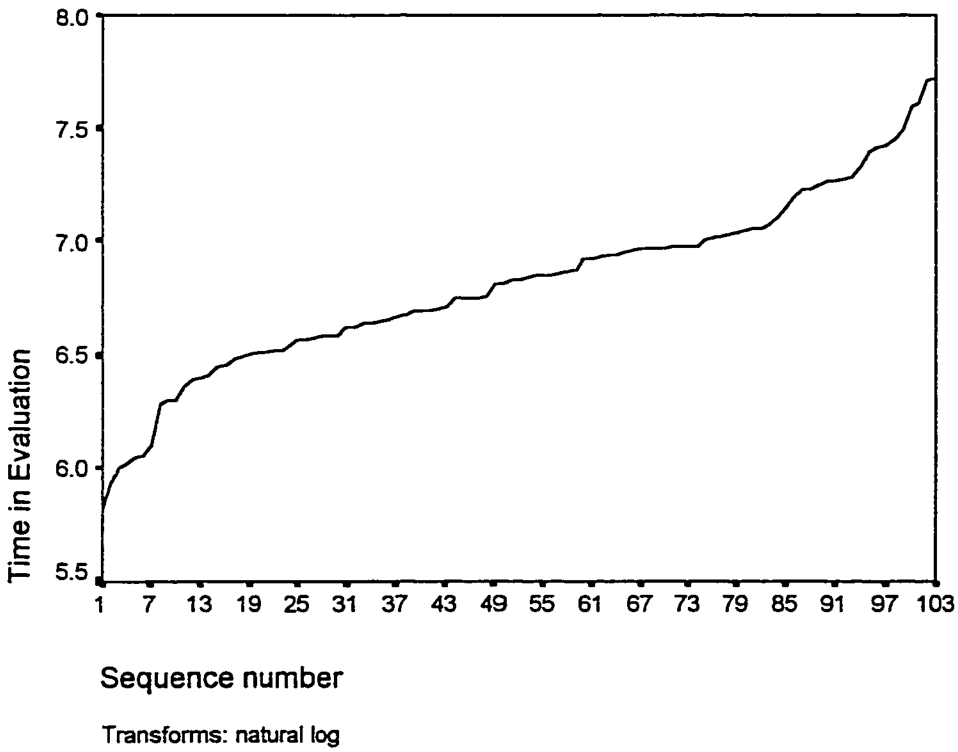


Figure I8: Plot of variable Time in Evaluation [Post transformation]

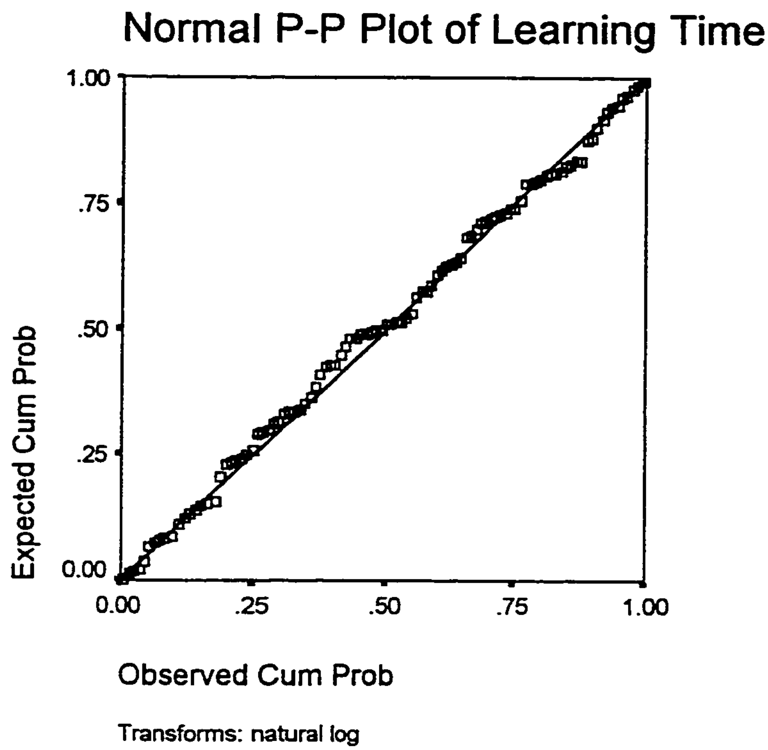


Figure I9: Normal P-P Plot Post Transformation

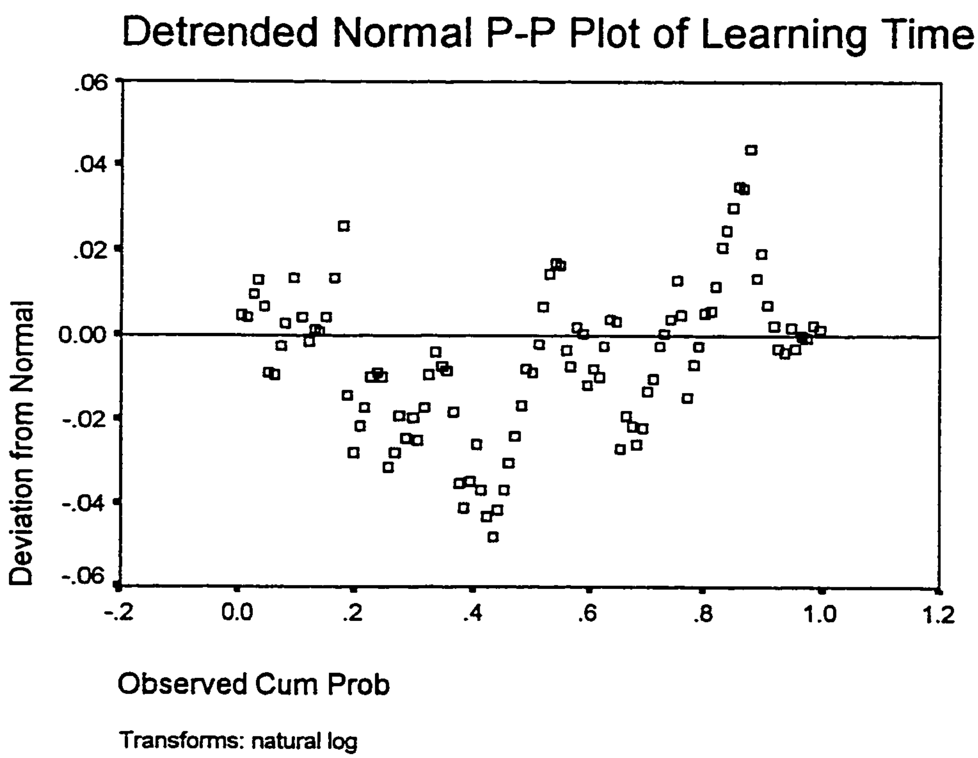


Figure I10: Detrended Normal P-P Plot Post Transformation

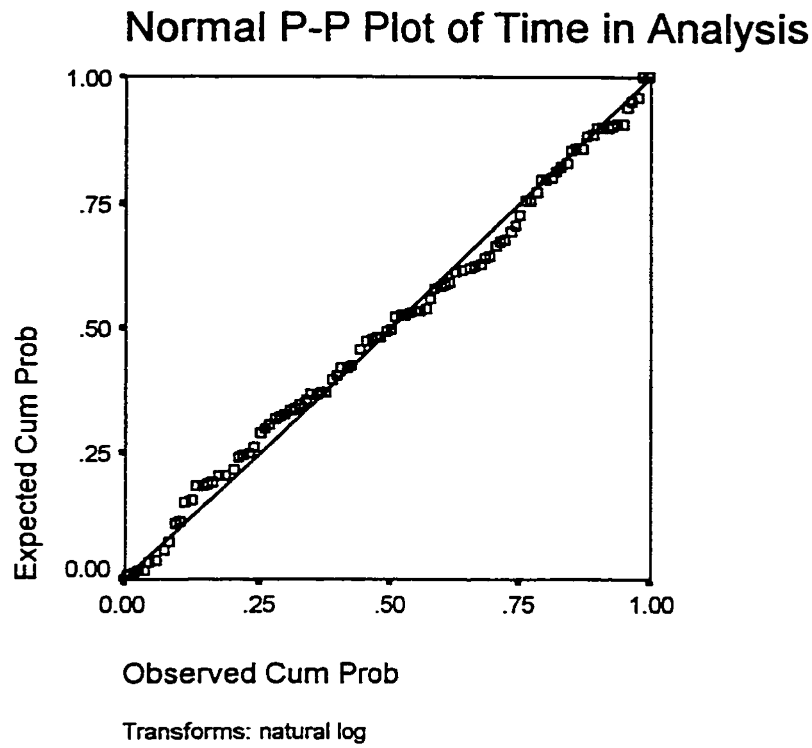


Figure I11: Normal P-P Plot Post Transformation

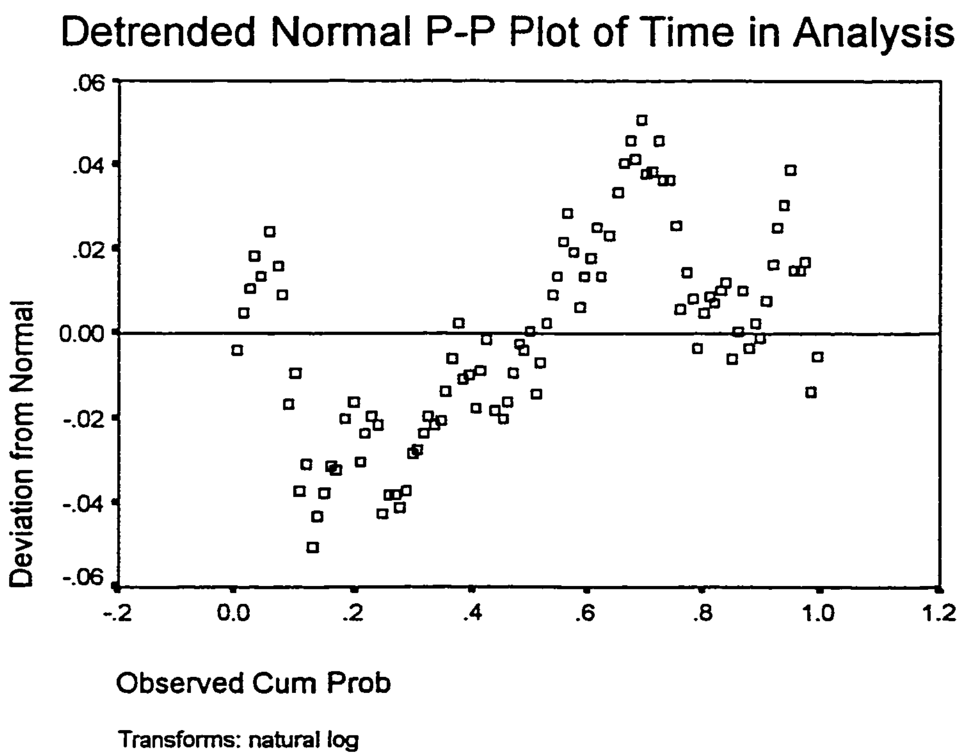


Figure I12: Detrended Normal P-P Plot Post Transformation

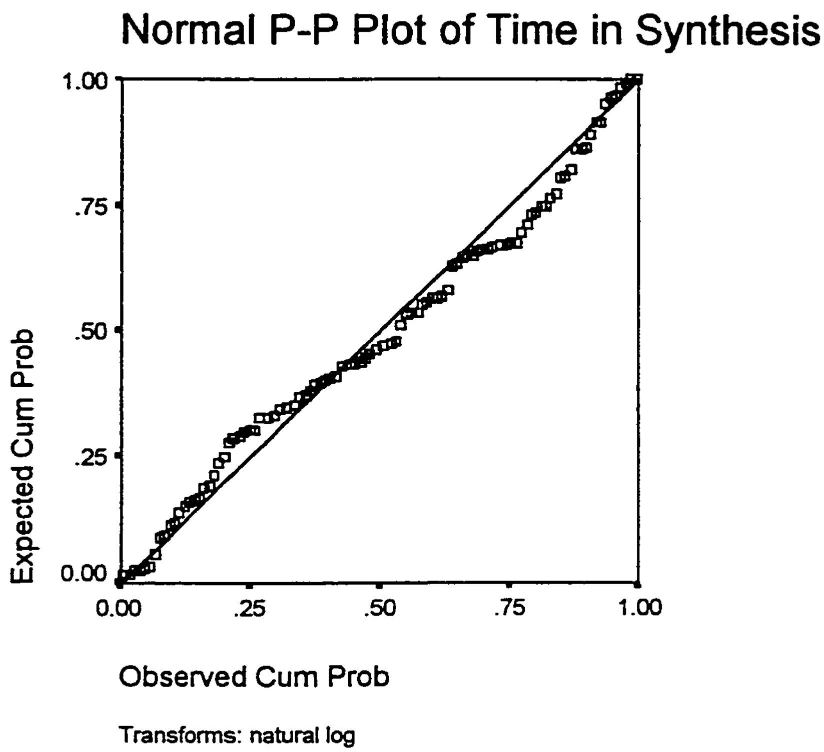


Figure I13: Normal P-P Plot Post Transformation

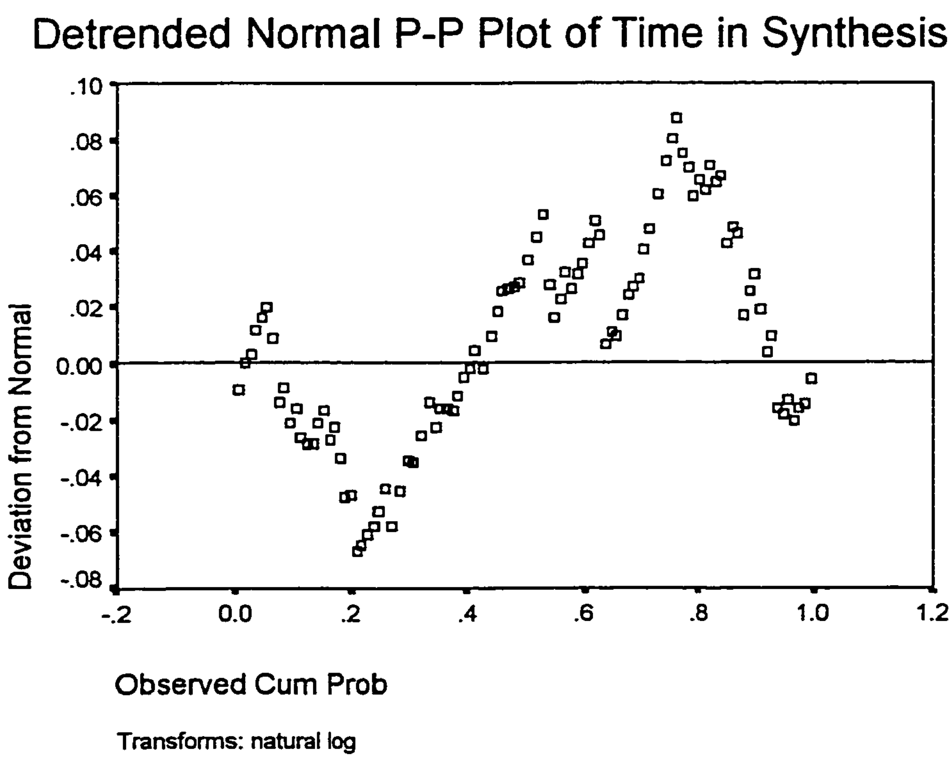


Figure I14: Detrended Normal P-P Plot Post Transformation

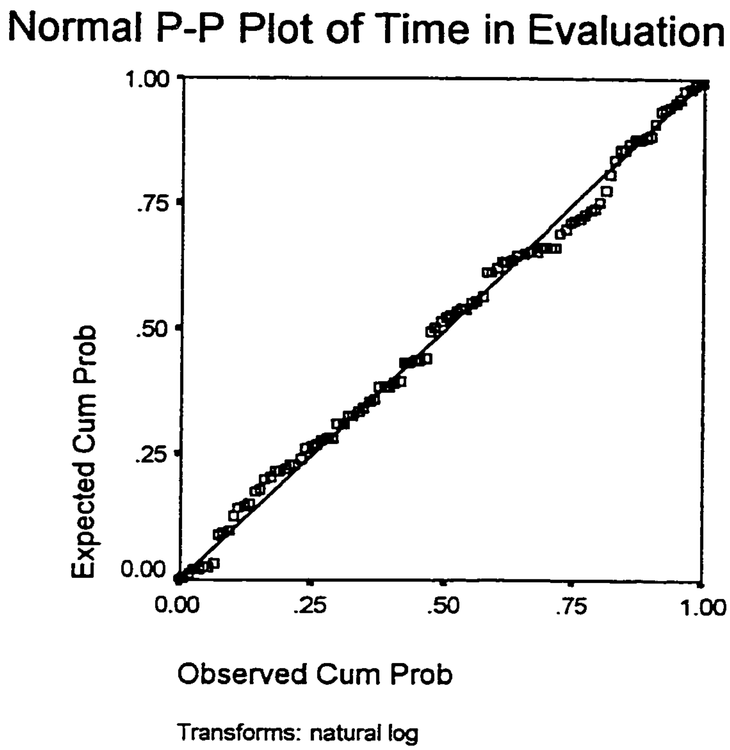


Figure I15: Normal P-P Plot Post Transformation

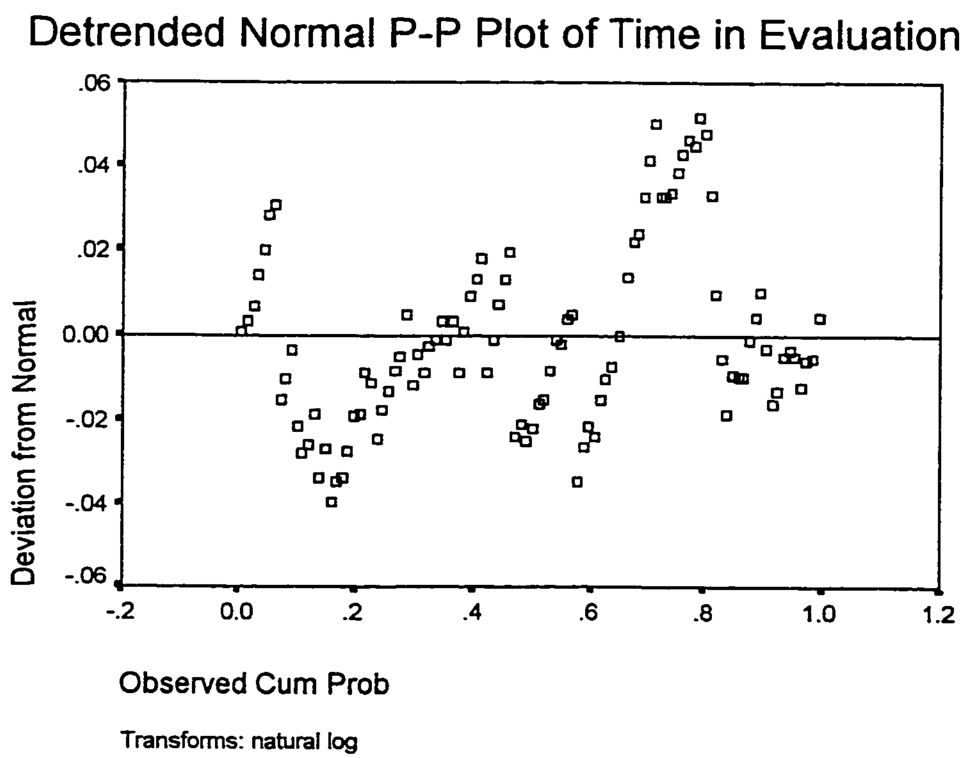


Figure I16: Detrended Normal P-P Plot Post Transformation

Appendix J

Correlation Matrix
between
Learning Time
and
the Dependent Variables
(Effectiveness and Efficiency
in
Analysis, Synthesis, and Evaluation)

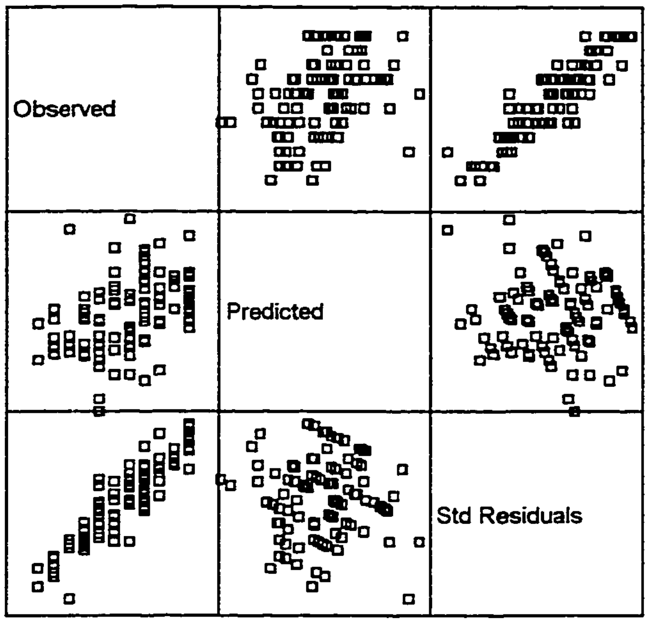
		Learning Time	Effectiveness			Efficiency		
			Analysis	Synthesis	Evaluation	Analysis	Synthesis	Evaluation
Learning Time		1						
Effectiveness	Analysis	0.340**	1					
	Synthesis	-0.022	0.028	1				
	Evaluation	0.241*	0.187	-0.109	1			
Efficiency	Analysis	0.288**	0.991**	0.026	0.173	1		
	Synthesis	-0.069	0	0.994**	-0.117	0.01	1	
	Evaluation	0.17	0.173	-0.119	0.990**	0.169	-0.12	1

Table I: Correlation Matrix, Significant at **α=0.01, *α=0.05**

Appendix K

Observed, Predicted, and Standardized Residuals Plots for the Dependent Variables

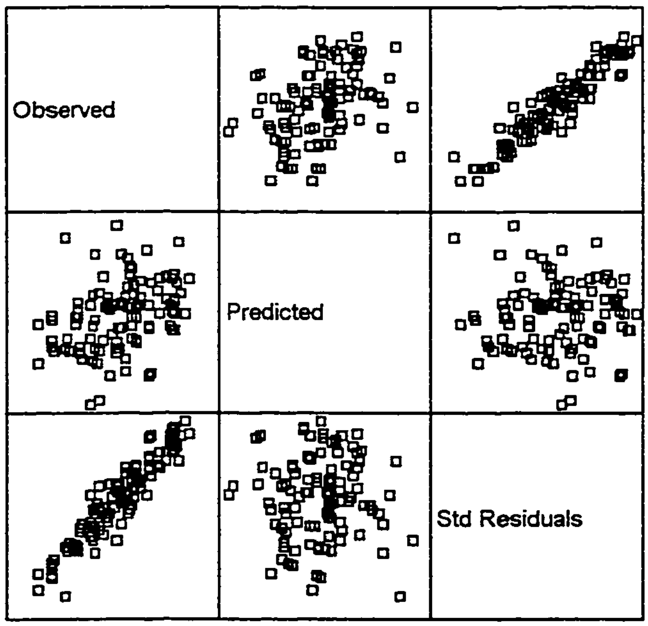
Dependent Variable: Effectiveness Analysis



Model:Intercept + LNTMOD + METHOD

Figure K1: Residuals Plots of Effectiveness in Analysis

Dependent Variable: Efficiency - Analysis



Model:Intercept + LNTMOD + METHOD

Figure K2: Residuals Plots of Efficiency in Analysis

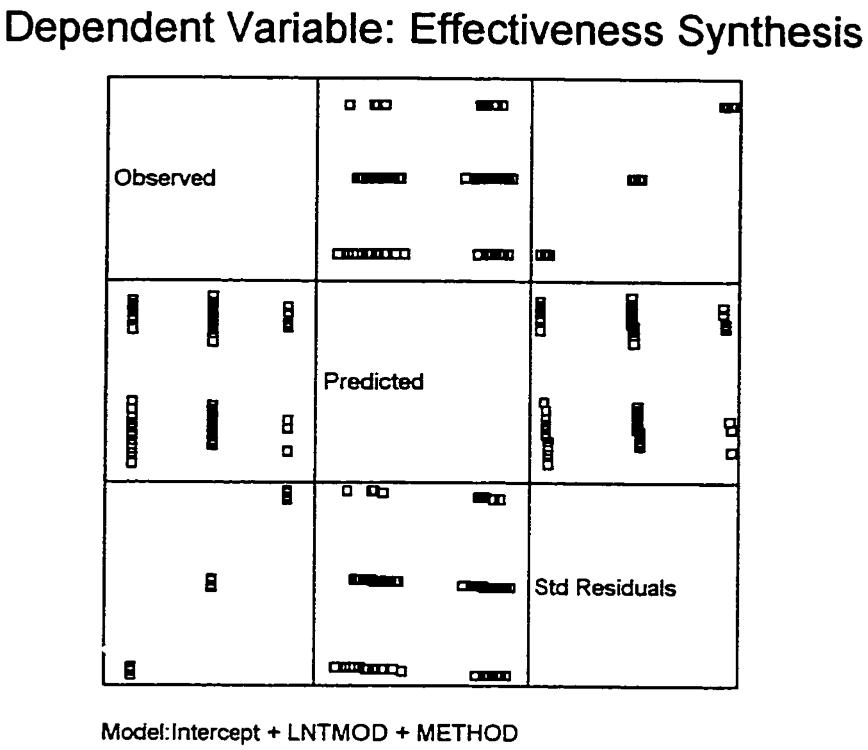


Figure K3: Residuals Plots of Effectiveness in Synthesis

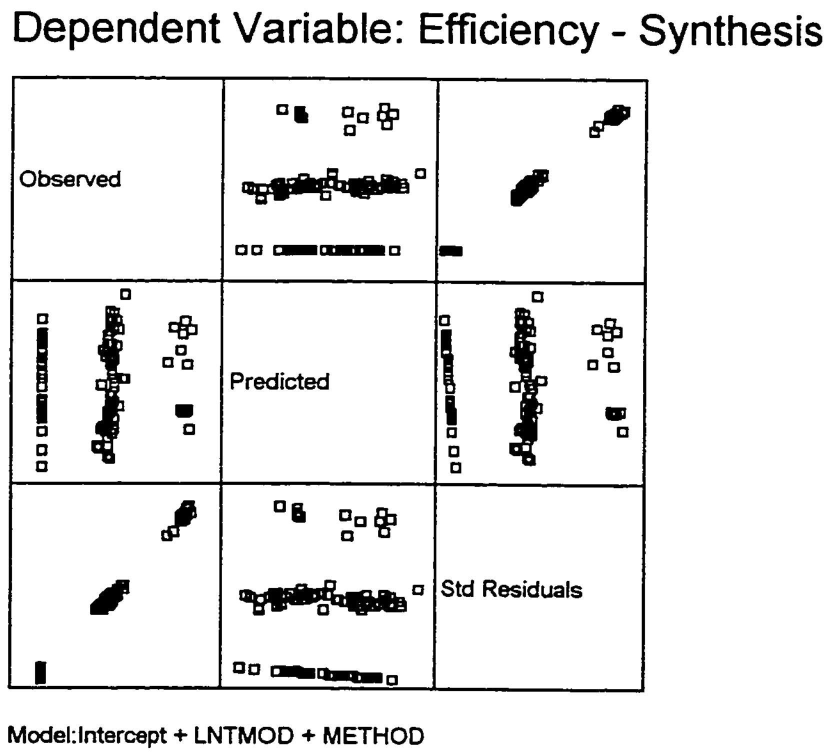


Figure K4: Residuals Plots of Efficiency in Synthesis

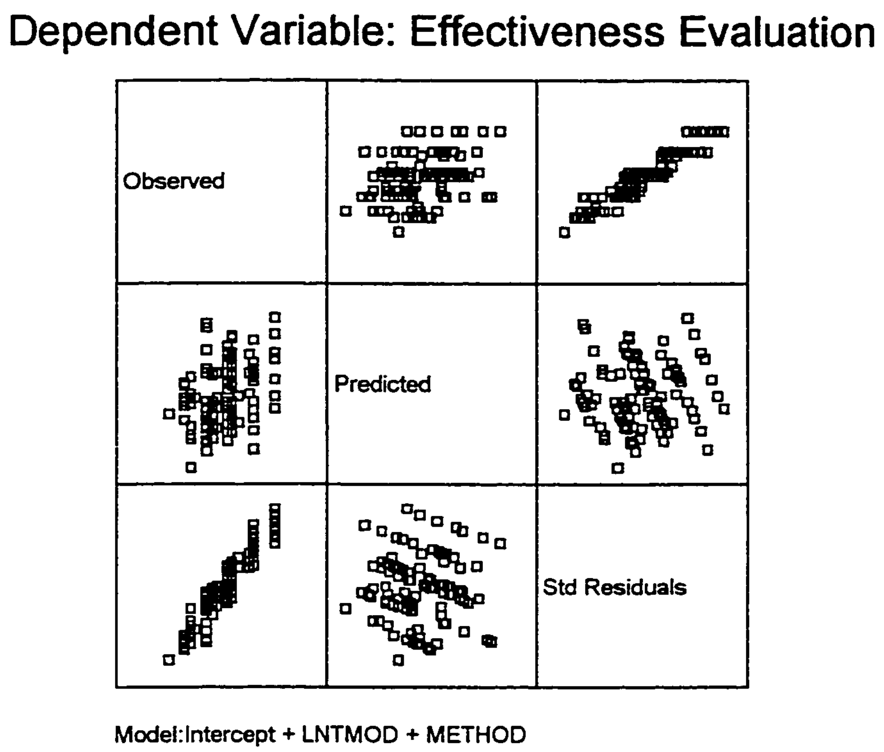


Figure K5: Residuals Plots of Effectiveness in Evaluation

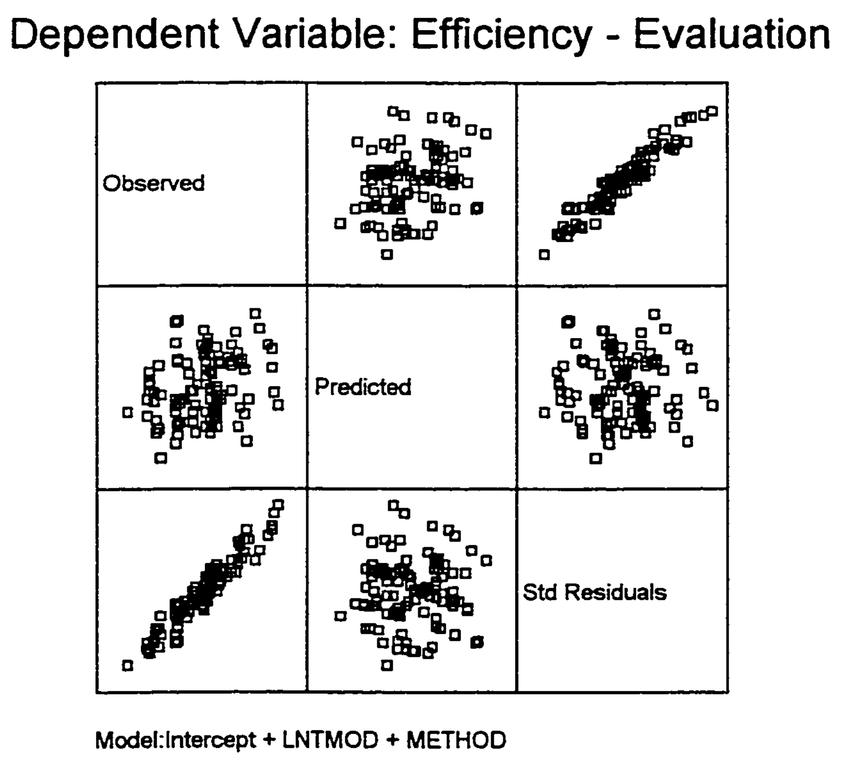


Figure K6: Residuals Plots of Efficiency in Evaluation

Appendix L

Correlation Matrix
between
Evaluation Factors
and
Total Score, Learning Time, Total Time and Hardware Use

		Score	Learning Time	Total Time	Hardware Use	Participation	Content: Motivation	Content: Usefulness	Content: Experience	Task: Motivation	Task: Satisfaction	Task: User friendliness	Test: Difficulty	Test: Challenge	Test: Motivation	Global Satisfaction
Score		1														
Learning Time		0.34	1													
Total Time		0.35	0.87	1												
Hardware Use		0	0	-0.1	1											
Participation		-0.1	0	0	0	1										
Content	Motivation	-0.3	-0.1	0	0	0.5	1									
	Usefulness	-0.2	-0.2	-0.2	0	0.3	0.6	1								
	Experience	-0.1	0	0	0	0	0.1	0.2	1							
Task	Motivation	0	-0.1	0	0.2	0.4	0.4	0.3	0.1	1						
	Satisfaction	-0.2	-0.3	-0.2	0.2	0.3	0.4	0.3	0.1	0.7	1					
	User Friend-liness	0	-0.1	-0.1	0.2	0.2	0.3	0.3	0.1	0.6	0.7	1				
Test	Difficulty	-0.2	-0.2	-0.2	0.2	0.2	0.3	0.2	0.4	0.3	0.3	0.2	1			
	Challenge	0	-0.2	-0.2	0	0.2	0.3	0.4	0	0.3	0.4	0.3	0	1		
	Motivation	-0.2	-0.3	-0.3	0.1	0.2	0.5	0.5	0	0.2	0.3	0.2	0.2	0.4	1	
Global Satisfaction		-0.2	-0.2	-0.3	0	0.4	0.5	0.5	0.1	0.4	0.6	0.4	0.2	0.4	0.5	1

Correlation Matrix, Significant at α=0.01, α=0.05

Appendix M

Histograms of Evaluation Factors

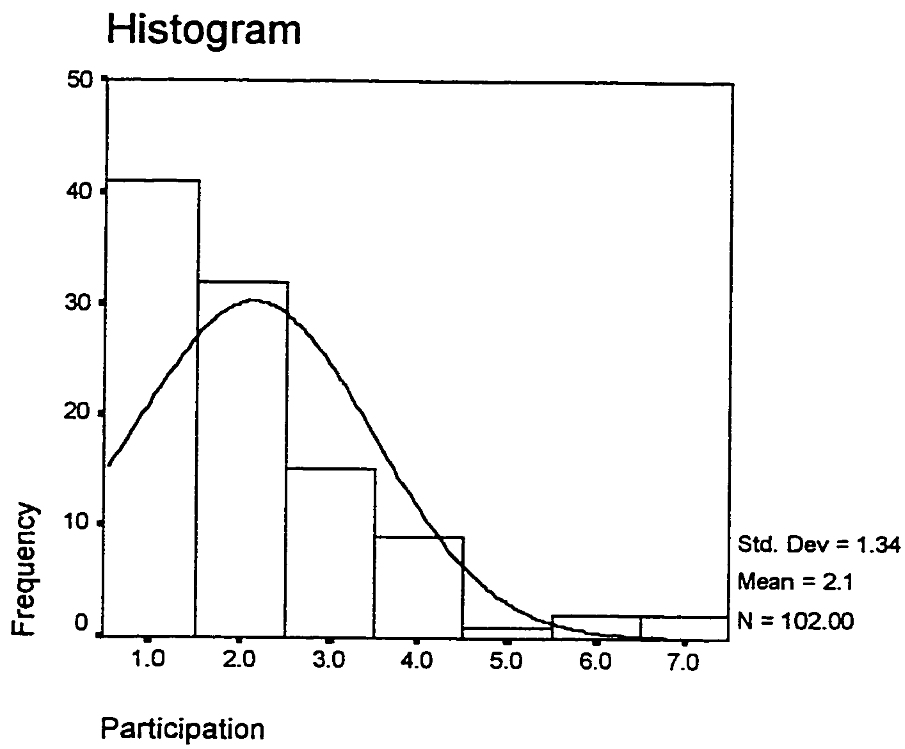


Figure M1: Histogram of Factor Participation

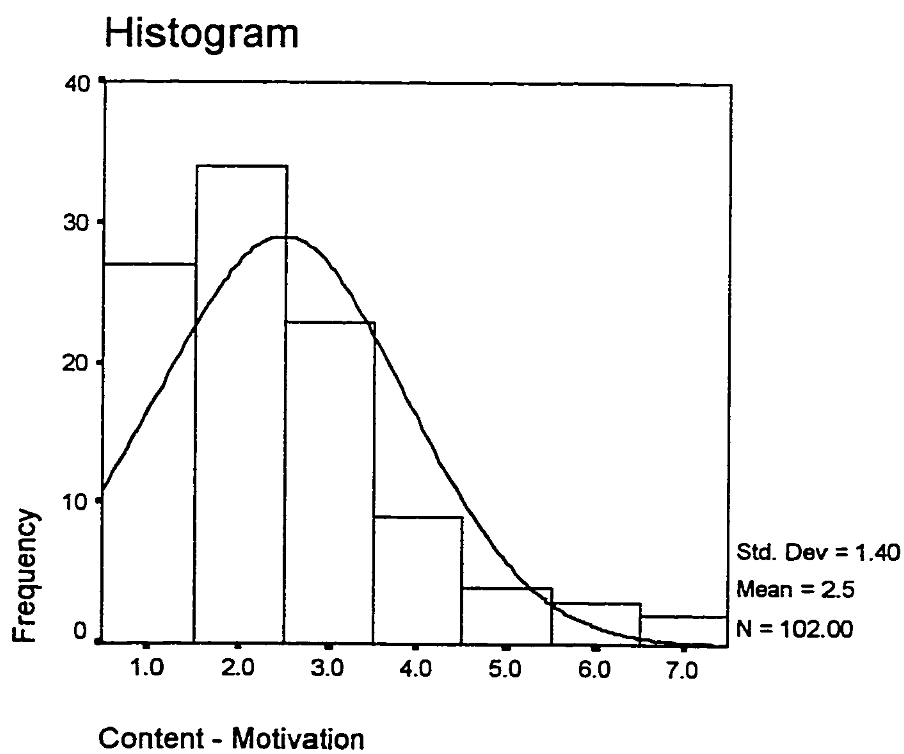


Figure M2: Histogram of Factor Content - Motivation

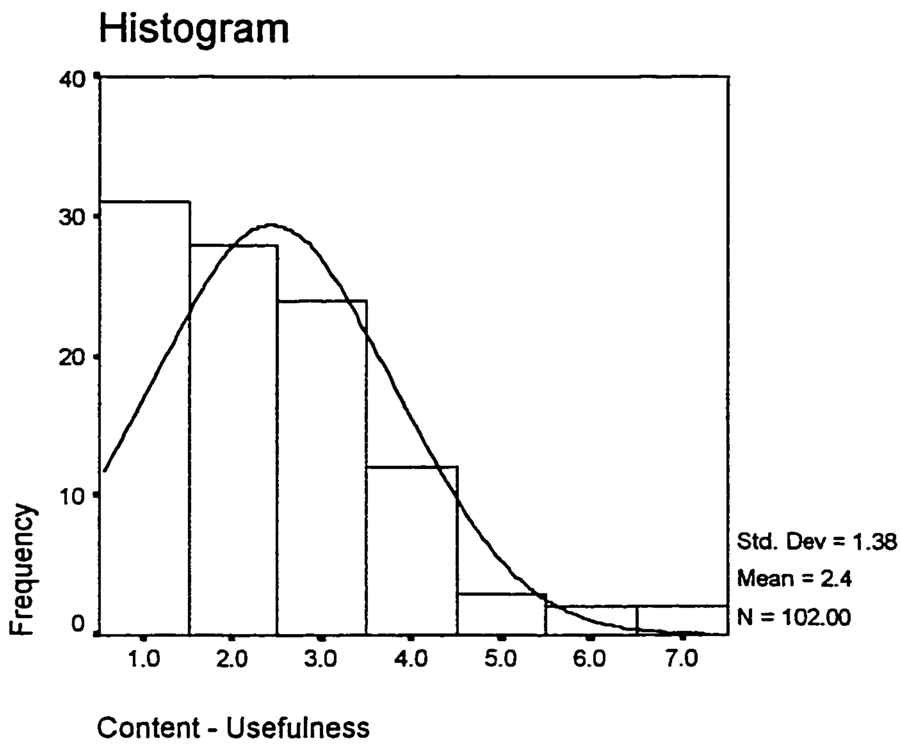


Figure M3: Histogram of Factor Content - Usefulness

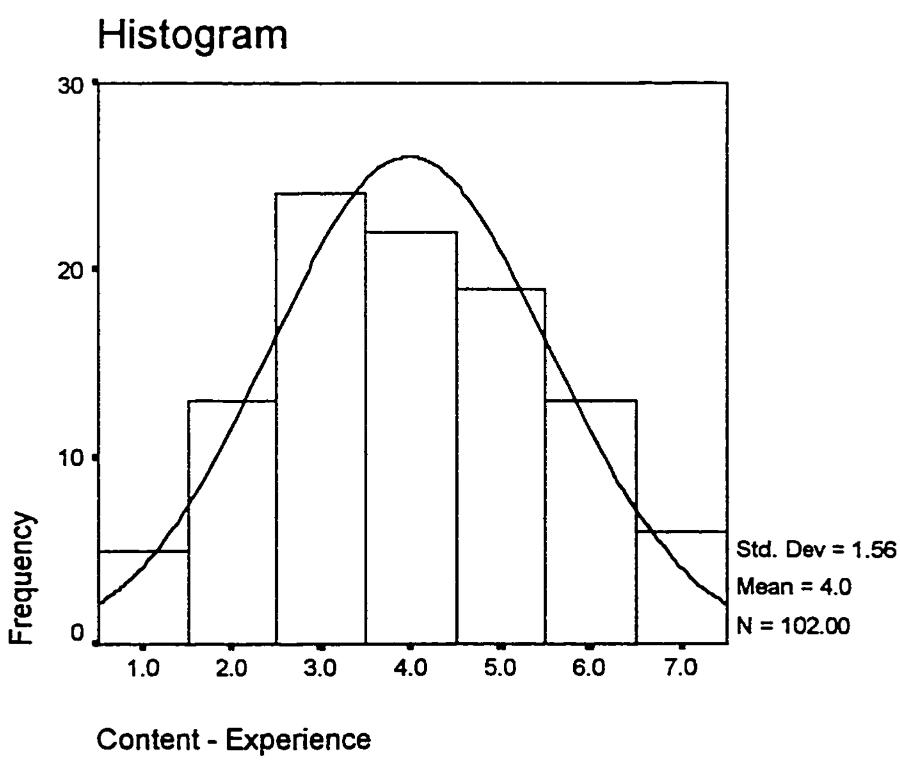


Figure M4: Histogram of Factor Content - Experience

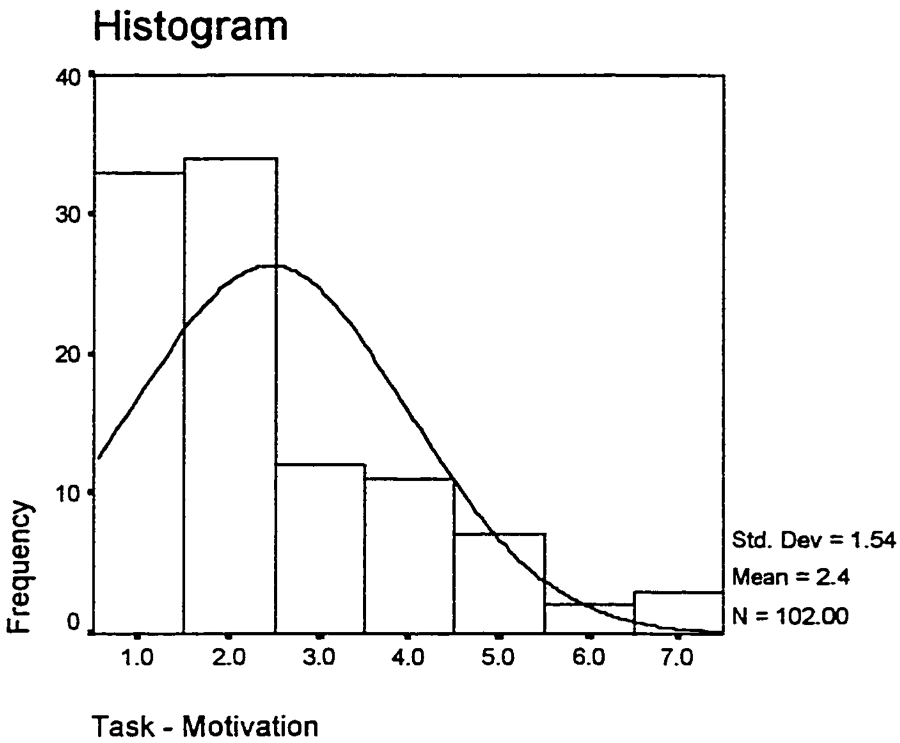


Figure M5: Histogram of Factor Task - Motivation

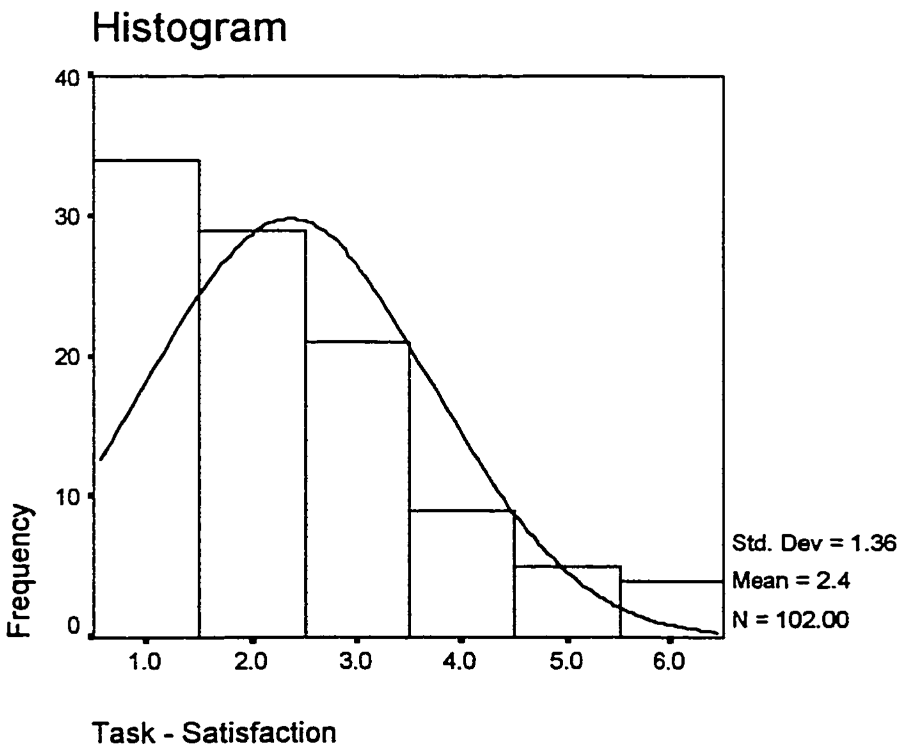


Figure M6: Histogram of Factor Task - Satisfaction

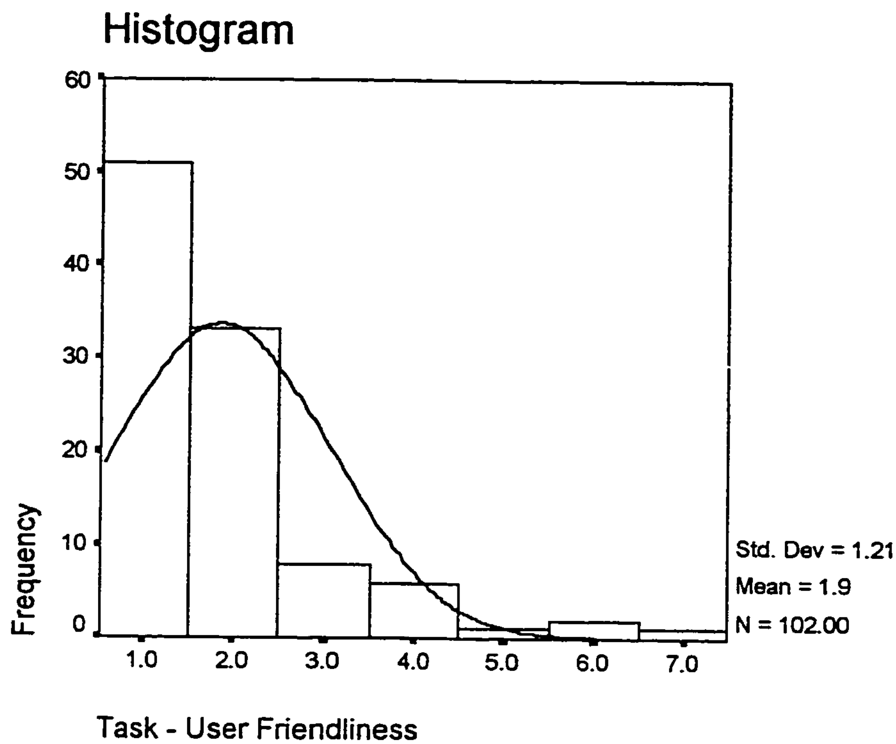


Figure M7: Histogram of Factor Task - User Friendliness

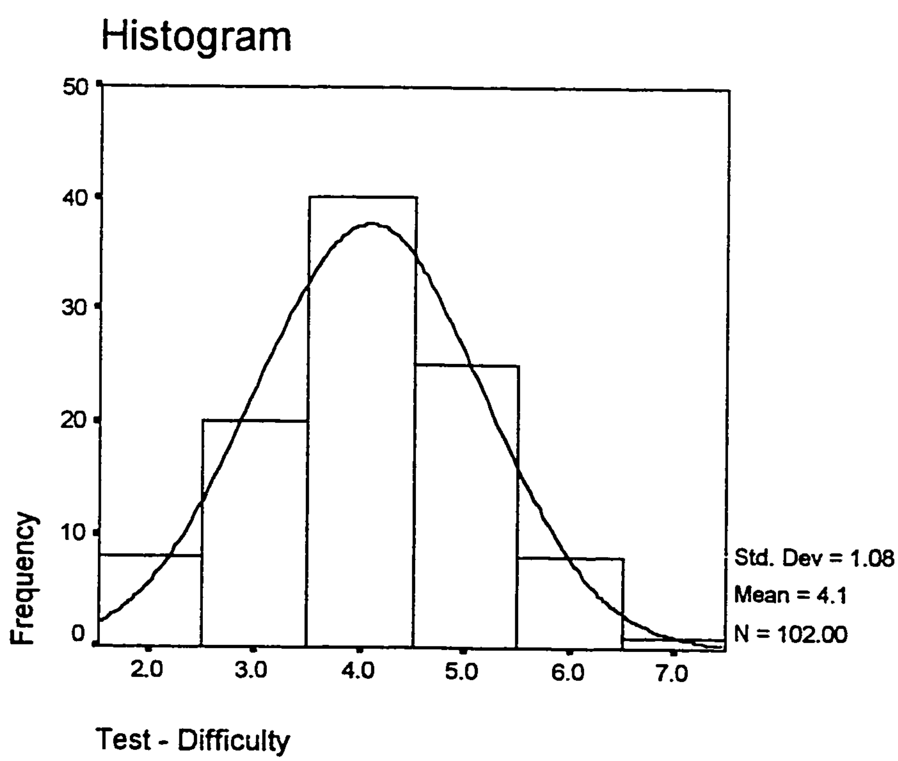


Figure M8: Histogram of Factor Test - Difficulty

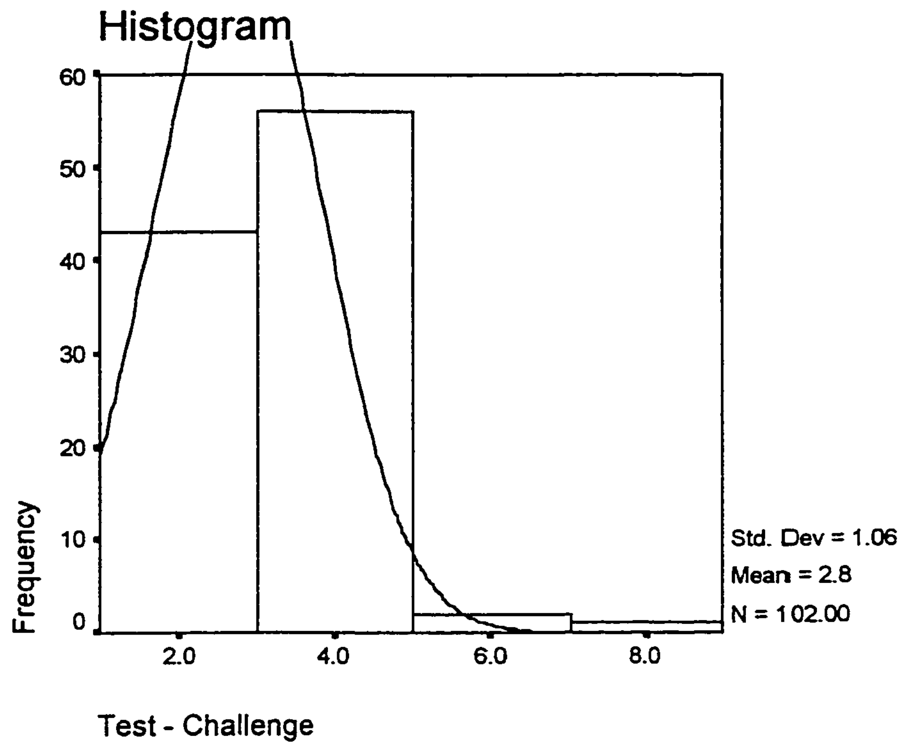


Figure M9: Histogram of Factor Test - Challenge

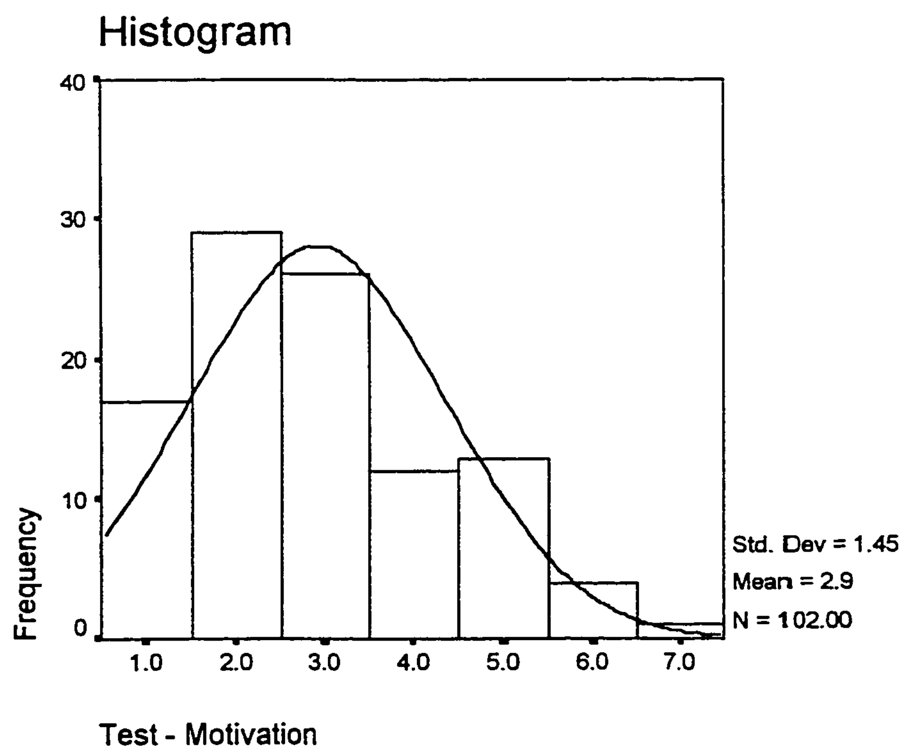


Figure M10: Histogram of Factor Test - Motivation

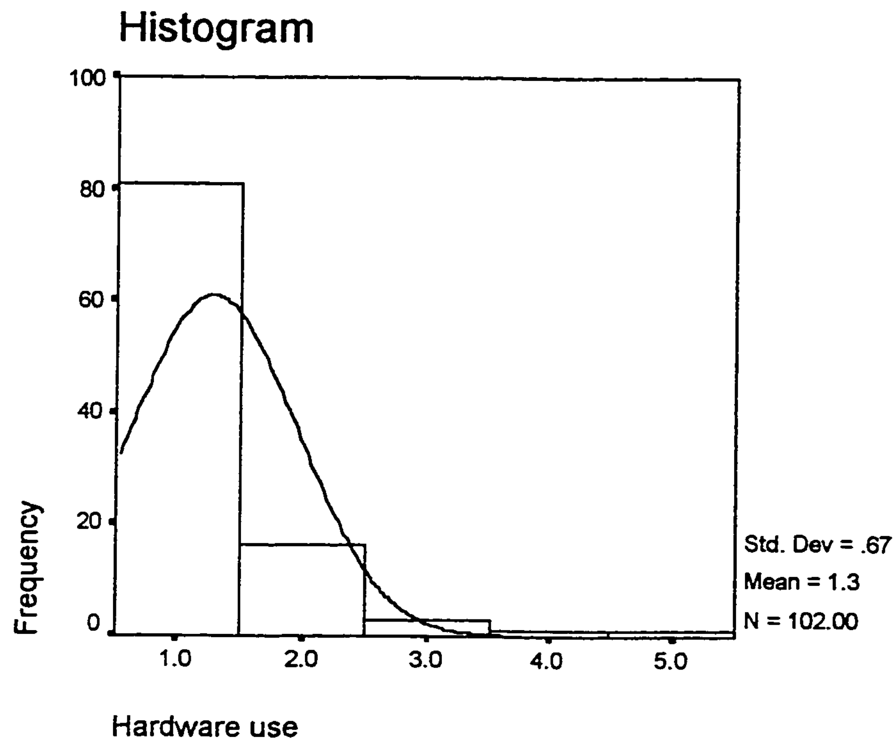


Figure M11: Histogram of Factor Hardware Use

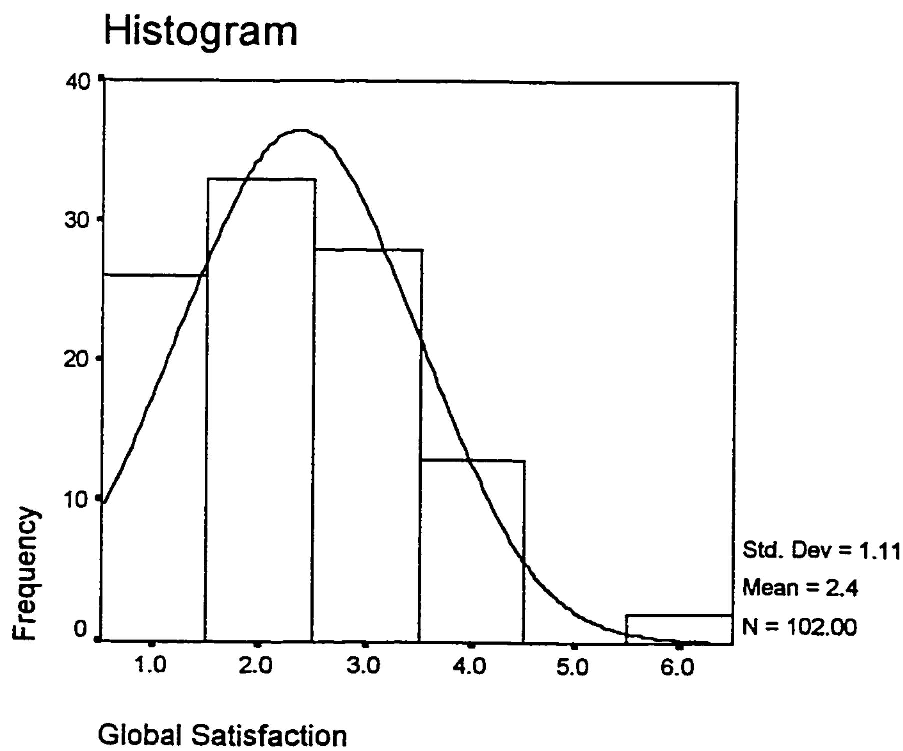


Figure M12: Histogram of Factor Global Satisfaction

Appendix N

Verbal Protocol Material

THINKING ALOUD PROTOCOLS

[To be read to participants in the Verbal Protocols Group]

When you are ready to proceed with the Test, please indicate so. At that time, we are going to turn on the tape recorder. What we need from you, while solving the test, is to try to think aloud. I guess you often do so when you are alone and working on a problem. I want you to say everything you happen to think of, no matter how irrelevant it may seem.

I am not primarily interested in your final solution, still less in your reaction time, but in your thinking behaviour, in all your attempts, in whatever comes to your mind, no matter whether it is a good or a less good idea or a question.

Read all information aloud, including question number and each one of the options.

Appendix O

Statistical Analysis
of
Verbal Protocol Data

Variable		N	Mean	Std. Dev.	Min.	Max
Learning Time		11	0.013	0.59722	0.3	54:24
Time	Analysis	11	0.3521	0.0938	0.1493	0.4972
	Synthesis	11	0.2361	0.11458	0.081	0.5035
	Evaluation	11	0.7035	0.35417	0.2854	37:10
Score		11	49.75	17.1	16.67	72.22
Effectiveness	Analysis	11	5.27	2.33	2	10
	Synthesis	11	2.45	1.81	0	6
	Evaluation	11	10.18	6.06	0	20
Efficiency	Analysis	11	46.73%	20.78%	17.85%	89.70%
	Synthesis	11	39.22%	28.77%	0%	93.94%
	Evaluation	11	47.50%	26.69%	0%	84.56%
Participation		11	1.91	1.14	1	4
Content	Motivation	11	2.64	1.21	1	5
	Usefulness	11	2.27	1.35	1	4
	Experience	11	2.91	1.3	1	5
Task	Motivation	11	2	1.18	1	4
	Satisfaction	11	1.45	0.69	1	3
	User Friendliness	11	1.36	0.92	1	4
Test	Difficulty	11	4.18	0.98	2	5
	Challenge	11	2.82	0.75	2	4
	Motivation	11	2.82	1.54	1	6
Hardware Use		11	1.09	0.3	1	2
Global Satisfaction		11	2.09	0.94	1	4

H_0: $\mu_{\text{Verbal Protocol}} = \mu_{\text{Without Verbal Protocol}}$
H_a: $\mu_{\text{Verbal Protocol}} \neq \mu_{\text{Without Verbal Protocol}}$

T-test for Equality of Means
(Equal Variances not assumed)

Variable		VP	n	Mean	Std. Dev.	t	df	Sig. (2-tailed)
Time	Analysis	0	92	0.3	0.12014	-1.67	14.21	0.116
		1	11	0.35	0.0938			
	Synthesis	0	92	0.24	0.13889	0.185	13.78	0.856
		1	11	0.24	0.11458			
	Evaluation	0	92	0.67	0.26111	-0.27	11.34	0.795
		1	11	0.7	0.35417			
Score		0	92	52.9	16.46	0.58	12.32	0.572
		1	11	49.7	17.09			

VP=1, Participated in Verbal Protocol Group.

In each case, we cannot reject the null hypothesis.