

**THE EFFECTS OF QUESTION CONSTRUCTION
ON EXPOSITORY TEXT COMPREHENSION**

**A Thesis
Submitted to the Faculty of Graduate Studies and Research
in Partial Fulfillment of the Requirements for the Degree
of Doctor of Philosophy
in the
Department of Educational Psychology and Special Education
University of Saskatchewan
Saskatoon, Saskatchewan**

By

Thelma M. Gunn

Fall, 2000

**The author claims copyright. Use shall not be made of the material contained
herein without proper acknowledgement.**



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-63952-5

Canada

In presenting this thesis in partial fulfillment of the requirements for the postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Request for permission to copy or to make other use of material in this thesis in whole or in part should be addressed to:

Head of the Department of Educational Psychology and Special Education
University of Saskatchewan
Saskatoon, Saskatchewan S7N 0X1

DEDICATION

This dissertation is
dedicated to my parents

G. Raymond and Audrey Barclay

ACKNOWLEDGEMENTS

Many people have supported and contributed to the completion of this dissertation. Of foremost importance was Dr. B.S. (Randy) Randhawa. His patience, guidance, and expertise have been appreciated more than words can say. Without him, the joys of research would have been less apparent. To my other committee members, Dr. Al Yackulic, Dr. Don Saklofske, Dr. Ivan Kelly, Dr. Jamie Campbell (Department of Psychology), and Dr. Walt Pawlovich (Chair of the Advisory Committee), I owe a great debt for their challenging insights and contributions. In addition, I would like to thank Dr. Linda Phillips (External Examiner) for her thought provoking questions and comments.

I am also indebted to Dr. Valerie Thompson (Department of Psychology) who enabled me to conduct research within the Department of Psychology. And finally, a special thanks to Monica Sliwa for her time and assistance as an interrater.

To my family, I owe an enormous debt. My daughter, Audrey, will one day be regaled with stories of how she nearly cost me my sanity, but at the same time restored it. A special thanks to my husband, Rob, who has endured hours of hand-wringing while I completed this dissertation. His support and optimism were much needed and appreciated.

Above all, I would like to thank my parents, Ray and Audrey Barclay. No two people could be more supportive and more proud of the accomplishments of their children. I do not know how any of this could have been completed without their unwavering encouragement. There are no words to express my gratitude for their love and support.

ABSTRACT

This study investigated three research questions: (1) In spite of text coherency and domain knowledge combinations, can we expect generic question stems to lead to greater measurable outcomes on memory (i.e., textbase model) and learning (i.e., situation model) variables as compared to signal words and/or unguided questioning techniques? (2) Will there be interaction effects on these variables suggesting that combinations of domain knowledge and text coherency are uniquely affected by generic question stems, signal words, or unguided questioning? (3) Will the employment of generic question stems lead to superior calibration of comprehension scores than those generated by signal words and/or unguided questioning conditions?

Sixty-three first year psychology students participated in the study. On the basis of their domain knowledge scores, which were ascertained during a screening procedure, participants were categorized as high or low domain knowledge. They were then randomly assigned to twelve treatment conditions in a 3 x 2 x 2 cross-factorial design. The first factor was treatment with three levels (i.e., generic question stems, signal words, and unguided questioning); the second factor was domain knowledge with two levels (i.e., high domain knowledge and low domain knowledge); and the third factor was text coherence with two levels (i.e., high text coherence and low text coherence).

Generic question stem and signal word participants were briefly trained to generate questions using their questioning strategies. Participants in the unguided questioning condition received no instructions. Upon reading the experimental and supplemental texts, all participants were required to generate, and respond to six questions. This was followed by a series of post-tests. Participants were administered four measures: a pre- and post-treatment sorting

task; a short answer post-test; a summary recall measure; and a calibration of comprehension task. All dependent variables derived from these measures were categorized as representing either a textbase model or a situation model. The experimental phase of this study was 120 minutes.

Results of this study suggest that high domain knowledge and structured questioning strategies are the most reliable predictors of text memory and text learning. Several main effects favouring high domain knowledge arose on text memory, and to a lesser extent, text learning measures. Significant results involving signal words and generic question stems were detected primarily through interaction effects and one main effect. While not always significantly superior, generic question stem participants outperformed their signal word and unguided questioning counterparts on the majority of text memory and text learning variables.

Based on the research findings, it appears that generic question stems engage the reader in a deeper level of processing. By making connections within the text and to one's prior knowledge, memory for text details and depth of processing increases. However, when instructional time is limited, as it was in this study, high domain knowledge is necessary for significantly superior memory of text and for specific structural connections.

TABLE OF CONTENTS

PERMISSION TO COPY	i
DEDICATION	ii
ACKNOWLEDGEMENTS	iii
ABSTRACT	iv
TABLE OF CONTENTS	vi
LIST OF TABLES	ix
CHAPTER ONE	
Introduction	1
Significance of the Study	2
Previous Text Processing and Questioning Research	13
Research Objectives and Hypotheses	14
Research Questions	16
Definition of Terms	17
CHAPTER TWO	
Review of the Literature	24
Constructivism	24
Historical Perspective	25
Situated Cognition and Cognitive Apprenticeship	27
Reciprocal Teaching	29
Reciprocal Peer Questioning	30
Text Comprehension	31
Measurement of Text Comprehension and	
Knowledge Acquisition	44
Expository Text versus Narrative Text	51
Expository Text and Questioning	54
Generic Question Stems	59
Calibration of Comprehension	65
Note-taking, Written Recording, and Memory	77
CHAPTER THREE	
Method	82
Participants	82
Research Design	83
Materials	84
Instruments and Analysis	87
Prior knowledge questionnaire	87

Post-test	88
Summary Recall	89
Sorting Task	91
Calibration of Comprehension	94
Procedure	95
Analysis of Data	99
 CHAPTER FOUR	
Results	101
Research Questions #1 and #2	101
Sorting data	102
Post-test	105
Summary recall	114
Inference analysis	114
Levels analysis	122
Research Question #3	126
Calibration of comprehension	128
Summary of results	130
 CHAPTER FIVE	
Discussion	135
Limitations of the Study	135
Conclusions Relative of the Research Questions	137
Question #1	137
Question #2	144
Question #3	151
Synthesis	153
Recommendations for educational practice and future research	160
 REFERENCES	163
 APPENDICES	190
A Constructivist Question Stems	190
B Domain Knowledge Questionnaire	193
C Training Materials	199
D Strategy Prompts	219
E Heart Disease Text (High Coherence)	223
F Heart Disease Text (Low Coherence)	227
G Blood Text	230
H Sorting Measure	234
I Calibration of Comprehension Measure	240
J Post-Test Measure	243
K Summary Recall Measure	249

L University Advisory Committee on Ethics in Behavioural Science Research Consent Letter	264
M Screening Procedure Consent Form and Study Description	266
N Experimental Procedure Consent Form and Study Description	269
O Question Response Recording Sheets	272
P Correlation Matrix for Question Type	275

LIST OF TABLES

Table	
1. Readability Scores	86
2. Inter-Rater Reliability Coefficients for Posttest	100
3. Inter-Rater Reliability Coefficients for Summary Recall	100
4. Means and Standard Deviations by Treatment, Domain Knowledge, Text Coherence, and Pre- and Post-Treatment Sorting Data	104
5. Combined Means and Standard Deviations for Treatment, Domain Knowledge, and Text Coherence: Independent Variables (Factors)	104
6. Means and Standard Deviations by Treatment, Domain Knowledge, Text Coherence, and Question Type	106
7. Combined Means and Standard Deviations for Treatment, Domain Knowledge, and Text Coherence: Independent Variables (Factors)	108
8. Tukey's Pairwise Comparisons for the Treatment x Text Coherence Combined Cell Means for the Post-Test Total Score	109
9. Treatment x Text Coherence Interaction: Significant Simple Effects for Univariate Analysis of Variance	112
10. Tukey's Pairwise Comparisons for the Treatment x Text Coherence Combined Cell Means for the Textbase and Elaborative Inference Scores	113
11. Means and Standard Deviations by Treatment, Domain Knowledge, Text Coherence, and Summary Recall Data	115
12. Combined Means and Standard Deviations for Treatment, Domain Knowledge, and Text Coherence: Independent Variables (Factors)	116
13. Treatment and Treatment, Text Coherence x Domain Knowledge Interaction: Significant Simple Effects for Univariate Analysis of Variance	119

14. Tukey's Pairwise Comparisons for the Treatment Combined Cell Means for the Elaborations and Reorderings	121
15. Means and Standard Deviations by Treatment, Domain Knowledge, Text Coherence, and Levels Analysis Data	123
16. Combined Means and Standard Deviations for Treatment, Domain Knowledge, and Text Coherence: Independent Variables (Factors)	124
17. Tukey's Pairwise Comparisons for the Treatment x Text Coherence x Domain Knowledge Combined Cell Means for the Level 1 Means of Treatments #2 and #3	127
18. Means and Standard Deviations for Domain Knowledge: Gamma and Pearson-Product Moment Correlation Coefficients	129
19. Means and Standard Deviations for Domain Knowledge: Confidence Rating and Post-Test Total Score	129
20. Summary of Results	132

CHAPTER ONE

INTRODUCTION

The proficiencies required to comprehend, remember, and learn from text are difficult to acquire. Each necessitates specific knowledge, skills, and strategies which are transmitted and obtained through extensive instruction and practice. In light of such inherent complexities, it is imperative that research expediting optimum text processing be pursued. This is especially true for expository text processing. To date, empirically-based instructional techniques and models of nonnarrative text have received little attention in the cognitive literature (e.g., Lorch, 1995; Lorch & van den Broek, 1997). Hence, any research furthering our understandings of nonnarrative text processing are well-warranted considering the contribution of expository text on classroom-based knowledge acquisition and utilization. Specifically, research involving expository text processing, learning, and memory is required for educational instructional development and application. Identifying the most effective means by which information can be drawn from text is salient when developing life-long learners.

A text comprehension technique which has received increased attention over the past two decades is questioning. It has been shown to solicit and cultivate metacognitive and self-regulatory knowledge, skills, and strategies by way of comprehension monitoring, creativity, increased attention, and higher level thinking, to name a few (e.g., Schank, 1986). One specific questioning

approach which has generated significant results is the generic question stem technique. However, this approach has typically been conducted in conjunction with lecture and/or lesson comprehension as well as in group settings. Moreover, it has been pursued almost exclusively by its creator, Alison King (e.g., 1989). Therefore, due to scant research findings and obvious genre limitations, further research combining generic question stems with expository text comprehension is warranted. Specific research questions worthy of investigation include whether guided generic question stems improve knowledge acquisition, retention, and transfer of coherent text and/or incoherent text at the micro- and macrolevels more so than signal word questions or unguided questions? Does a reader's level of domain knowledge in combination with coherent and/or incoherent text at the micro- and macrolevels affect performance outcomes? Do guided generic question stems enable the reader to better calibrate expository text comprehension more so than signal word questions or unguided questions? Thus, there is much to be learned so that educators may improve learners' acquisition, utilization, and maintenance of text-derived knowledge and comprehension skills and strategies.

Significance of the Study

The ability to read and comprehend text is a well-recognized goal of instructional practice. Research findings indicate that word recognition processes must be accurate and automatized in order for text comprehension to occur (e.g., Adams, 1990). More simply, text comprehension is contingent upon the basic decoding skills of the reader. Fortunately, a great deal of laudable research concerning orthographic, morphemic, and phonologic processing has been conducted enabling a clearer understanding of the

complexities involved in word processing (e.g., Adams, 1990; Stanovich, 1986, 1989). Such research has enabled the development of instructional models targeting adaptive reading development as well as reading recovery.

Word recognition, however, is only one of many reading comprehension processes. Mental representations of text, text coherence, and topic familiarity all contribute to comprehensibility (Beck, McKeown, Sinatra, & Loxterman, 1991; Just & Carpenter, 1987; Spires & Donley, 1998; van Dijk & Kintsch, 1983). Despite a reader's technical proficiency, his or her ability to comprehend a piece of text may be impeded if one or more of these components are deficient. For instance, a mental representation of text is comprised of a textbase and a situation model. The textbase model encompasses those elements and relations which are directly derived from the text itself (McNamara & Kintsch, 1996). For adaptive development of this model, the reader must have syntactic and semantic knowledge of the text as well as be able to infer its microstructure and macrostructure. The situation model involves meaning-making. Simply, it is a representation of the situation depicted in the text (i.e., goals, events, actions, characters) (Zwaan, 1996). It is constructed using the textbase as well as the reader's prior knowledge and experience. Various sources contribute to the building of a situation model including knowledge about language, the world, the specific communicative situation, as well as personal experiences (McNamara & Kintsch, 1996). These sources help to "transform what by itself is only an isolated memory structure into something that relates to, and is integrated with, the reader's personal store of knowledge and experience" (McNamara & Kintsch, 1996, p. 252). It is therefore apparent that failures in generating a mental representation can seriously undermine comprehensibility as well as understanding of the text.

Also integral to text comprehension is text coherence or text structure (e.g., Kintsch, 1994a). The manner in which a text is written and/or structured either aids or impedes its comprehensibility. The microstructure of text is comprised of local text properties (i.e., connectives, argument overlap, pronominal reference). These properties give the text coherence at a propositional or local level (McNamara & Kintsch, 1996). Text macrostructure refers to the global organization of the text (i.e., topic headers and topic sentences). Because macrostructure specifies the most important portions of the text, it helps the reader during the construction of a strong mental representation.

According to Kintsch (e.g., 1985), if either the micro- or the macrostructure is poor, development of an effective textbase model may be curtailed. Its absence in the comprehension process is further exacerbated when coupled with low domain knowledge (i.e., topic familiarity). That is, a reader who enters the reading context with a great deal of domain knowledge may be equipped to build a strong situation model, despite the absence of a coherent textbase. The superiority of the situation model is forged when the reader is qualified to fill-in informational gaps and discrepancies with prior knowledge. Consequently, retrieved memory networks are connected to that which is intelligible within the text thereby creating an adequate understanding of the text. Such a deep level of processing can override otherwise damaging comprehension difficulties. If the text lacks coherency at both local and global levels and the reader's domain knowledge is insufficient, textbase and situational models are less likely to arise (Kintsch, 1985). Whether this condition can be remedied with a method of comprehension scaffolding is worthy of consideration.

When reading narrative text, it is typical for the reader to develop an adaptive mental representation. In most instances, the reader is equipped to build both textbase and situation models. That is, the text is typically structured in a manner which is familiar to the reader. Its contents also closely correspond with common experiences by describing everyday events, personality traits, human emotions, and the pursuit of goals (e.g., Graesser, Bertus, & Magliano, 1995; Graesser & Kreuz, 1993; van Dijk & Kintsch, 1983). Even if the text is poorly structured at either local or global levels, as may be the case in some literary attempts (e.g., Zwaan, 1996), there is likely to be sufficient prior knowledge for the reader to generate a situation model. Needless to say, this is not always the case with expository text. Expository text is laden with technical jargon, theoretical underpinnings, and implicit assumptions regarding previous knowledge. These preconditions are further exacerbated when authors of descriptive text expect readers to fill-in these gaps and/or to overlook poor micro- and macrostructure (Graesser et al., 1995). To demonstrate the frequency and subsequent damage generated by such a belief, Britton, van Dusen, Gulgoz, and Glynn (1989) revised 52 instructional texts which fell short of criteria embodied in a standard readability formula. These formulas convert quantitative information about word frequency and syntactic complexity (i.e., .90 correlation of word length and sentence length) into grade level equivalents. This information was then used to either select or revise textbooks for specific grades. When compared to original text retention scores, 3 out of 5 text experts significantly improved recall and recognition scores on immediate and delayed retention tests by making revisions to text readability. Britton et al. (1989) argue that these results demonstrate that readable texts increase comprehensibility and information retention.

Britton et al. (1989) also note the importance of text signaling. Text signals are integral components for adaptive global coherence or macrostructure (Kintsch, 1998). The original instructional texts, which concern Army job tasks, general science, philosophy, and history, were all revised to accommodate the signaling hypothesis which states that signals that visually delineate the structures of complex texts assist in the processing and retention of text. This position was verified in the results of their Experiment 1 indicating that changes in signaling are correlated with improvements in retention.

Due to the vast number of instructional texts used in their study, there is evidence to suggest that a surprising number of textbooks are poorly structured at the local and global level. Beck and McKeown (1989), who define coherence as "the extent to which the sequencing or ordering of ideas in a text makes sense and the extent to which the language used in discussing those ideas makes the nature of the ideas and their relationships apparent" (p. 50), found in their examination of expository selections coherency disruptions within introductory passages as well as within internal passages which digressed to new topics. Because there are no inherent structural properties in expository text, as there are in narrative text, Beck and McKeown state that authors typically create their own structures for communicative effect. Inherent narrative structures contain elements of a goal or problem, attempts to attain the goal or solve the problem, and subsequent resolutions. This overall organization, or macrostructure, is familiar to the narrative reader thereby assisting in text comprehensibility. When macrostructure is vague and unclear, as is often the case in expository text, it is difficult to build a mental model of the situation; a model vital for learning and not just remembering text content (Albrecht, O'Brien, Mason, & Myers, 1995; Albrecht & O'Brien, 1993; Beck, McKeown, &

Gromoll, 1989; Tapiero & Denhiere, 1995). As Beck and McKeown (1989) discovered, this is additionally hindered when the text lacks local coherency. That is, poor microlevel organization can ultimately lead to macroprocessing difficulties. Similar findings concerning poor local and global coherence arose in Beck, McKeown, Sinatra, and Loxterman's (1991) review of a fifth grade social studies text. Upon revision of these coherency breaks, students were able to recall more material and correctly answer more questions than those who read the original revision. Therefore, it appears that poor structure is common within expository text materials and may be the cause of many comprehension difficulties.

Construction of generalized meaning, or macrostructure, has been shown to be especially difficult for adolescents who have shown a developmental lag in macroprocessing of expository text (E. Kintsch, 1990). As suggested, this can only be exacerbated by instructional text deficient in local and global coherence. In most cases, a writer or speaker conveys a message which is coherent. When it is lacking, the reader can remedy local coherency breaks by making connections either between propositions in short-term memory or through easily accessible pieces of general knowledge. Local coherency can also be built by way of salient connections to earlier parts of the text (McKoon & Ratcliff, 1995). Nevertheless, in spite of their intrinsic and spontaneous nature, local coherency building strategies can fail. Proficient reading is a carefully orchestrated compromise between speed and accuracy. Due to this tradeoff, not all information is checked for accuracy. The results of such breakdowns have been evidenced by way of the Moses illusion (e.g., Kamas & Reder, 1995; Reder & Kusbit, 1991) as well as Barton and Sanford's (1993) example of burying the survivors of a plane crash. Hence, even

proficient readers and those with high domain knowledge can fail to make important coherency links through poor inference judgments as well as inappropriate semantic, associative, and pragmatic connections to other concepts (Kamas & Reder, 1995; McKoon & Ratcliff, 1992; 1995). These coherency failures are especially detrimental for those with low domain knowledge and/or poor comprehension skills.

Despite the potential for coherency breakdowns amongst proficient, knowledgeable readers, there is some evidence to suggest that high domain knowledge coupled with poor macrostructure and/or poor local coherence produces the largest learning and memory gains (e.g., E. Kintsch, 1990; McNamara & Kintsch, 1996; McNamara, Kintsch, Songer, & Kintsch, 1996). It has been theorized that low coherence text incites knowledgeable readers to retrieve the vital information and knowledge necessary to rectify deficient causal connections, argument repetitions, topicality links, and so forth. Simply put, the reader engages in a deeper level of text processing. It appears that this does not occur for low domain knowledge students. Because they are lacking the requisite domain knowledge, low knowledge students must be helped in maximally benefiting from poorly structured text. McNamara and Kintsch (1996) and McNamara et al., (1996) demonstrated that low domain knowledge subjects perform best when using high coherence text. However, as indicated above, it is not always possible to provide students such materials since many texts are written with poor microstructure and/or poor macrostructure. With limited educational funds and resources, it is near impossible to tailor textbooks to the needs of all students within a class. So what can be done to help all readers, with either high and low domain knowledge, generate adaptive textbase and situation models when they are asked to read high or low

coherence text? Furthermore, what can be done to help them determine whether they have accurately comprehended the text?

Much of the comprehension process involves the degree to which a reader monitors his or her reading behaviour (Weaver, Bryant, & Burns, 1995). Educators typically assume that experienced readers accurately judge whether they have successfully or incorrectly comprehended a piece of text. This has been defined as "calibration of comprehension" or the correlation between what the reader believes he or she comprehended and what was actually retained (e.g., Glenberg & Epstein, 1985, 1987; Glenberg, Sanocki, Epstein, & Morris, 1987; Maki & Berry, 1984; Maki, Foley, Kajer, Thompson, & Willert, 1990; Weaver, 1990). Surprisingly, research results indicate that correlations are typically close to zero (Glenberg, Wilkinson, & Epstein, 1982; Glenberg & Epstein, 1985, 1987; Glenberg, Sanocki et al., 1987; Morris, 1990). Readers often fail to reliably monitor their reading processes. However, several studies have demonstrated that the inducement of additional processing can improve calibration accuracy (Glenberg, Sanocki et al., 1987; Maki et al., 1990; Maki & Swett, 1987). Techniques include: a) making predictions regarding the memorability of contradictory information (Maki & Swett, 1987); b) reading and filling-in deleted letters of words versus reading intact text (Maki et al., 1990); and c) inducing self-generated feedback concerning one's level of text comprehension (Glenberg, Sanocki et al., 1987).

In regards to self-generated feedback, Glenberg, Sanocki et al. (1987) suggest pre-testing as a feedback technique. Pre-test results provide information concerning comprehension/retention correlations. These results are believed to alert the reader to areas of difficulty within the text thereby leading to improved calibration scores on the post-test. More simply, the reader

uses the results from a pre-test to generate conclusions regarding what was comprehended and retained. The expectation is that the reader will then return to the original text and remedy comprehension/retention discrepancies.

While pre-testing has remedial potential, it is both impractical (Glenberg, Sanocki et al., 1987) and ill-equipped to develop strategic awareness. Pre-tests do not directly instruct the reader how to monitor comprehension or how to learn from text, they simply provide results concerning perceived comprehension/retention correlations. Due to such fallibility, Glenberg, Sanocki et al. (1987) suggest that advanced organizers and textual signals are more appropriate. These techniques help the reader to monitor his or her own comprehension/retention of text during the reading process. Moreover, they are believed to connect the text in such a manner as to induce a clear cognitive representation. As indicated earlier, textual signals are integral to the development of macrostructure. Moreover, Mannes and Kintsch (1987) conducted a study using advance organizers and their effects on text memory and learning. Two types of organizers were tested: one which was congruent with the target text and the other being incongruent. Those students in the incongruent-advance organizer condition showed superior performance on inference verifications and on difficult creative problem-solving tests. Those students in the congruent-advance organizer condition performed better than their counterparts on cued-recall and recognition tasks. These performance measures are indicative of situation model development and textbase model development, respectively. Therefore, like the studies cited above concerning levels of domain knowledge and text coherence, it appears that additional processing generates a deeper, more connected understanding of text information and the text situation. While tailoring materials to the students

and/or to the text content appears to have merit both in this study and in the studies mentioned above, it does invoke the same arguments regarding educational funds and resources. Finding a solution which can be employed with all students and used with any type of text must be sought and tested. Any approach which successfully blends individualized tailoring and group administration and instruction is highly advantageous. Ideally, the approach would induce readers to engage in personalized levels of cognitive processing despite its generic structure or format. One such approach is questioning. That is, questioning techniques can be administered to large groups, they can be used individually or cooperatively, and they can be personalized.

Alison King (e.g., 1989) is the primary author and proponent of Reciprocal Peer Questioning and generic question stems. She has successfully demonstrated the utility of this approach with adult populations, and on occasion with adolescents in lecture/lesson comprehension settings. The approach involves the development of questions using generic stems. The stems are not so much guidelines but structures upon which questions are constructed. Their purpose is to invoke the construction of internal and external connections within the material being studied and the student's prior knowledge. Upon question creation, students are expected to discuss their questions within small groups.

There is a substantial body of research supporting the merits of King's approach in regards to performance outcomes. Her work was, however, restricted to lecture/lesson comprehension and cooperative learning. Whether this approach would be advantageous with expository text and individualized question generation and response can only be inferred. Like lecture/lesson presentations, expository text can be loosely structured and speculative of

learner knowledge and comprehension skill. As previously mentioned, it can also be incoherent making comprehension difficult; particularly for those with low domain knowledge. By employing generic question stems, coherency may be easily achievable at both the local and global levels making the construction of textbase and situation models feasible.

So how might generic question stems help a reader to create appropriate micro- and/or macrostructures as well as appropriate textbase and situation models? The answer may be as simple as devising stems which require the reader to make links within the material and to his or her domain knowledge. While there is evidence that readers automatically process text micro- and macrostructure, even in the absence of task demands, the outcomes are not always accurate. Thus, it may be necessary to have some means in which to monitor their formation. For example, Otero and Kintsch (1992) demonstrated that in the presence of a strong text macrostructure, readers tend to ignore contradictory material because of the macrostructure's overarching strength. Had the readers been required to carefully construct a textbase and situational understanding, illogical macrostructure development may be detected. Moreover, the question stems may require the readers to engage in local and global coherency building strategies such as developing questions which invoke argument repetitions, bridging inferences, and causal connections. These too would stimulate the reader to detect contradictions within the text, to fill-in possible information gaps, and above all, to develop global understanding of the situation and the content of the text.

Generic question stems could have their strongest impact on textbase and situation model development. The stems could be designed to induce textbase question generation, which would require information from a single

sentence of the text, as well as situation question generation, which would necessitate bridging inferences, elaborative inferences, and problem solving. Moreover, as suggested by Schmalhofer (1996), in order to generate a question, the situation model must be searched for some insufficiently tested hypothesis. Instructing learners to generate their own questions would not only necessitate the conscious development of a situation model, it would ensure that a complete model was formulated.

In conclusion, this study was predicated upon the hypothesis that textbase and situation model development and monitoring would enable readers to better calibrate their comprehension of text. By way of self-questioning, the correlation between what is comprehended and what is retained would be improved. Without strategic explicitness and scaffolding, the acquisition and utilization of adaptive knowledge, skills, and strategies was hypothesized to be unachievable. According to previous research, such explicitness was especially vital for those with high domain knowledge and low domain knowledge as they have been shown to be either overly confident or lacking in coherency building strategies, respectively .

Previous Text Processing and Questioning Research

As previously mentioned, generic question stems have not been studied in conjunction with low and/or high domain knowledge and ill-structured expository text. To this author's knowledge, the only researcher to study generic question stems is King (e.g., 1989) and her research has been restricted to lecture format and/or classroom lessons as well as cooperative learning environments. While there was some consistency amongst the studies in controlling for prior knowledge, it rarely warranted special consideration

(King, 1990a, 1991a, 1991b). When it was mentioned, the purpose was to acknowledge and subsequently avoid additional effects on knowledge acquisition (King, 1989, 1992b, 1994a; King & Rosenshine, 1993).

In regards to Kintsch and his associates' (e.g., Kintsch & van Dijk, 1978) work on text processing and comprehension, they have not employed questioning as a means to incite more active processing by the reader, to develop textbase and situation models, or as a means to enhance local and global coherency.

Research Objectives and Hypotheses.

Employing expository text to study the utility of generic question stems was warranted. Because generic question stems have already been shown to exhort an understanding of lecture and lesson based material, it was hypothesized that they would continue to incite a deep processing of information even when it was presented in a complex written form (i.e., poor microstructure and poor macrostructure). They would also cause the reader to build strong textbase and situation models by constructing a representation of the text and linking it with prior domain knowledge. These suppositions were expected to manifest by way of performance scores.

As previously discussed, well-structured text makes fewer demands on the reader's comprehension skills, memory, and prior knowledge. However, because there are fewer demands, some readers may process material at a shallower level. This is especially the case for those students with high domain knowledge (McNamara & Kintsch, 1996; McNamara et al., 1996). For students with low domain knowledge, text must be explicit and well-structured. If it is not, those lacking the required knowledge will find it difficult to build either a

textbase or situation model. A hypothesized solution for both high and low domain knowledge students who are reading either well-structured or ill-structured text was generic question stems. It was anticipated that stems would expedite active processing as well as the creation of much needed text structure by promulgating connections both within the material (textbase model) as well as to previously learned material (situation model).

The utilization of expository text, as compared to narrative text, poses additional reading challenges which were addressed. Expository text is often laden with terminology and information which the reader may find unfamiliar or difficult to process. Without specific and effective inducements (i.e., generic question stems), text content may be dismissed or marginally processed.

Finally, generic question stems were hypothesized to induce comprehension awareness thereby leading to improved calibration of comprehension scores. Because the question stems are explicit and because they require additional processing for completion purposes, they were expected to better indicate comprehension hits and misses. Thus, the reader would be better able to gauge his or her own comprehension levels.

It is important to mention that the generic question stems devised by King were modified by this author. The modifications helped to make the stems more specific and utilitarian (see Appendix A). Moreover, they were coded suggesting their appropriateness for textbase and situation model development. It is important to note that King did not code her generic question stems. She did, however, typically ask questions on her post tests requiring students to go beyond the presented material by a making inferences, analyzing information, elaborating on ideas, and applying it in new contexts. In this study, generic question stems were not only be coded, but all performance variables were also

categorized as textbase variables or situation variables. This meant that outcomes could be definitively categorized as memory or learning outcomes.

In summary, an experiment was conducted to ascertain whether generic question stems significantly aid in the comprehension of expository text (locally incoherent and globally incoherent/locally coherent and globally coherent) when combined with either topic familiarity or topic unfamiliarity. Signal word questioning and untrained questioning techniques served as comparison conditions. Independent variables included treatment, text coherency, and domain knowledge. Dependent variables included summary recall inference (i.e., propositions, generalizations, elaborations, and reorderings) and levels analysis (i.e., Level 1, Level 2, Level 3, and Level 4); post-test performance (i.e., textbase, elaborative inferencing, bridging inferencing, and problem solving); sorting scores (i.e., pre- to post-treatment); and calibration of comprehension (i.e., calibration correlations and confidence ratings). Both the summary recall and post-test were based upon recall as recognition tests tend to be less sensitive to weaknesses in a text's coherence (Estes, 1995; Weaver, Bryant, & Burns, 1995). Furthermore, the calibration of comprehension measure was given to determine whether generic question stems assist participants in accurately gauging text comprehension and whether they increase reader confidence. It was hypothesized that the generic question stem conditions would not only significantly outperform the other conditions on all variables, calibration of comprehension would also be more accurate.

Research Questions.

This study was predicated upon three research questions. They are as follows:

(1) In spite of text coherency and domain knowledge combinations, can we

expect generic question stems to lead to greater measurable outcomes on memory (i.e., textbase model) and learning (i.e., situation model) variables as compared to signal words and/or unguided questioning techniques?

(2) Will there be interaction effects on these variables suggesting that combinations of domain knowledge and text coherency are uniquely affected by generic question stems, signal words, or unguided questioning?

(3) Will the employment of generic questions stems lead to superior calibration of comprehension scores than those generated by signal words and/or unguided questioning conditions?

Definition of Terms

Expository Text

Expository text is written to expose information or ideas. Similar to narrative text, it may be used for entertainment, persuasion, or aesthetic purposes. However, unlike its entertainment counterpart, its primary purpose is to inform the reader (Graesser & Goodman, 1985).

Text Microstructure

The local structure of the text. Microstructure is comprised of complex propositions and their interrelationships (Kintsch, 1998). Thus, microstructure is the sentence-by-sentence, propositional information within the text.

Text Macrostructure

Derived from the microstructure through mapping rules (e.g., deletion, generalization, and construction), macrostructure is the hierarchically ordered set of propositions which represent the global structure of the text. It is a

process of reduction which characterizes the process of abstracting or summarizing a text (van Dijk, 1995). It consists of macropropositions (i.e., the main points) and their hierarchical relations. Macrostructure is sometimes explicitly signaled by organization signals (e.g., headings, topical overviews, topical summaries) (Lorch & Lorch, 1996), but is most often inferred by the reader (Kintsch, 1998).

Proposition

The basic composite unit of language. Often analyzed in terms of a predicate and one or more arguments thereby characterizing the internal semantic structure of a sentence (McNamara et al., 1996; van Dijk & Kintsch, 1983). Predicates that take arguments in text analysis are adverbs, adjectives, and quantifiers (Perfetti & Britt, 1995). Complex propositions are schematic representations of the meaning of sentences. They are comprised of a predicate slot and a circumstance slot, specifying the time and place of the action or state. Each of these may be modified (McNamara et al., 1996).

Coherence

Coherence is when sequences of clauses and sentence sequences intuitively "hang together" (van Dijk, 1995). It is the connectedness of the text. Coherence at the local level is accomplished by carrying propositions over from one processing unit or cycle to the next, or from one sentence to the next (generally between three sentences) (e.g., Graesser, Bertus, & Magliano, 1995; Kintsch, 1994a; O'Brien, 1995). Nouns and pronoun arguments in the current clause are linked with explicit arguments and propositions in the previous text. Also, connections (i.e., and, or, because, so) link adjacent clauses and clauses.

There are also causal chains and connections, bridging inferences and argument repetition (McKoon & Ratcliff, 1992; van den Broek, Risdén, & Husebye-Hartmann, 1995). Thus, local coherence between sentences has some direction or continuity.

Coherence at the global level is enhanced by the appropriate reinstatement in the textbase of propositions from a preceding part of the text. It also involves the organization of local chunks of information into higher order chunks and the linking of an incoming clause to a clause which occurred much earlier in the text and is no longer available in active memory (e.g., Graesser, Bertus, & Magliano, 1995; Kintsch, 1994a; O'Brien, 1995). Therefore, it is somewhat dependent upon local coherence and requires more globally oriented bridging inferences, causal connections, and in particular, argument repetition (Tapiero & Denhiere, 1995). It is global coherence which gives the text overall unity.

Both local coherence and global coherence are frequently referred to as a good microstructure and good macrostructure.

Reading Comprehension

The capacity to understand and make sense of the intention and meaning of written text. Product comprehension indicates data which the reader is able to or willing to report. It may take the form of explicit factual, text-based information. Process comprehension indicates how the reader makes sense of, considers, and constructs text both during and after reading engagement (Goodman, 1984).

Textbase Model/Representation

A stable copy of the textual information which contains concepts and meanings of the text as well as a microstructure and macrostructure (Mannes & St. George, 1996). Textbase construction involves the propositionalization of text contents so that a reader may develop a sense of coherence at local and global levels. To develop a coherent, memorable representation of the textbase, both the writer and the reader must activate the most important propositions so that they are frequently processed and recently stored in memory. A strong textbase will result in greater recall and text memory (Kintsch, 1994a).

Situation Model/Representation

It is not solely textbased. It is a combination of prior knowledge, including the goals and attitudes of the reader, and text information. A situation model is a mental representation of the people, actions, setting, and events that are explicitly mentioned or inferentially suggested (Graesser & Zwaan, 1995). Reflects a successful integration of new and old information. Useful in situations which require problem-solving or divergent thinking. Textbase is good for recall and verbatim representations which contain purely textbase information whereas the situation model contains information regarding what the text is about (Mannes & St. George, 1996; McNamara & Kintsch, 1996).

Mental Representation

A two stage connectionist algorithmic process beginning with the construction of networks, or modules of activated concepts found within text. This is followed by the integration process which involves the strengthening of

compatible units or concepts and the weakening of those which are incompatible. This process continues until a stable state of mind is achieved; that is, the reader's mental representation (Kintsch, 1988, 1994a, 1994b).

Prior Knowledge

Noted by Alexander, Schallert, & Hare (1991) and Dochy (1996), prior knowledge is often poorly defined or misused. In response, they define prior knowledge as the whole of a person's knowledge. Thus, "prior knowledge is dynamic in nature"; "available before a certain learning task"; "is structured"; "can exist in multiple states (i.e., declarative, procedural, and conditional knowledge)"; "is both explicit and tacit in nature"; "and contains conceptual and metacognitive knowledge components" (Dochy, 1996, p. 229).

Domain knowledge

Domain knowledge represents a body of knowledge which is specific to a subject or realm of study. As such, domain knowledge requires prior knowledge and is unlikely to transfer from one domain to another without explicit transfer inducing instruction (Dochy, 1996). To be characterized as having "low" domain knowledge, one should only receive a score of 10 to 25% on a well-constructed domain-specific questionnaire/test (Voss & Bisanz, 1985). Those "high" in domain knowledge should receive an inverse score of 75 to 90%. However, for the purposes of this study, domain knowledge was classified by way of a median-split procedure. With a sample size of 80 participants, domain knowledge scores were too homogeneous for Voss and Bisanz's (1985) grouping suggestion. The median of the distribution of 80 scores was 20.9; while the mean was 24.7 and the standard deviation was 14.2.

The mean of the high domain knowledge participants was 37 (S.D. = 9.0) while the mean of the low domain knowledge participants was 12.4 (S.D. = 4.3).

Calibration of Comprehension

The correlation between an individual's comprehension confidence rating concerning a read passage and his or her performance on subsequent comprehension questions (Glenberg & Epstein, 1985, 1987; Glenberg, Sanocki, Epstein, & Morris, 1987). Also referred to as "metamemory for text" (Maki & Berry, 1984; Maki & Swett, 1987), this form of comprehension monitoring correlates one's subjective assessment of knowledge gained from reading and his or her performance on an objective test (Glenberg et al., 1987). High calibration indicates an appropriate assessment of one's state of knowledge upon reading a passage. For example, a well-calibrated individual may be aware of his or her inability to comprehend what was read, confident of this lack of understanding/acquisition, and subsequently demonstrative of this lack of knowledge on a comprehension post-test.

Low calibration suggests either an underconfidence or an overconfidence in one's state of knowledge in relation to what was actually comprehended. Simply, the amount of knowledge an individual believes to have understood/acquired is either greater than or less what was actually understood/acquired. For instance, a poorly-calibrated subject may claim to have learned a substantial amount from what was read when in fact he or she misunderstood or failed to comprehend the gist of the passage.

Generic Question Stems

Generic question stems are partially completed sentences which cue the subject to create internal connections and external connections (Mayer, 1989, 1992) by way of question generation.

ex. How is _____ associated with, or related to what we have learned/read _____ before?

Are _____ and _____ related in any way? Explain.

Internal and External Connections

Similar to selective combination and selective comparison (Sternberg, 1985), respectively, internal connections involve the organization of selected information from the presented material into a coherent whole while external connections link some or all of the newly acquired information to other prior knowledge structures (Mayer, 1989, 1992).

Transfer

A demonstration that what has been learned is generalizable and useful. Manifestations include finding and mapping higher-level analogies, noticing problem similarity and abstracting a schema, and mapping productions from a source to a target problem (Kotovsky & Fallside, 1989).

CHAPTER TWO

REVIEW OF LITERATURE

This chapter is organized around several bodies of literature which support the theoretical and experimental underpinnings of this study. They include: a) constructivism and its impact on memory and learning; b) text comprehension research, including Kintsch's (e.g., van Dijk & Kintsch, 1983) construction-integration model (CI) ; c) measurement of text comprehension and knowledge acquisition; d) differences in the comprehension of expository text versus narrative text; e) expository text comprehension research using questioning; f) generic question stem research; g) research involving calibration of comprehension; and h) the importance of note-taking, written recording, and memory.

Constructivism

Over the past decade, there has been sufficient evidence to suggest that the acquisition, utilization, and maintenance of knowledge, skills, and strategies are enhanced by way of elaborative (e.g., McDaniel & Donnelly, 1996; Woloshyn, Paivio, & Pressley, 1994) and generative processing (e.g., Kourilsky & Wittrock, 1992; Wittrock, 1990). In support of their philosophic stance, constructivists have argued both explicitly and implicitly that elaboration and generation are the cornerstones of constructivism (e.g. Cobb, 1988; Pressley, Harris, & Marks, 1992). In spite of the number of factions within the

constructivist camp (i.e. radical constructivism, cognitive constructivism, social constructivism), each with a slightly different theoretical inclination, all constructivists concur that learners must activate and engage prior knowledge to ensure that learning is both meaningful and memorable. Thus, it behooves educational researchers to consider notable constructivist models when developing learning theories and instructional programs.

Historical Perspective

Constructivism is rooted in the writings of Jean Piaget (1970). He asserted that humans are active knowledge constructors who constantly seek a state of equilibrium between the mind and the environment. To Piaget, such a pursuit is innately-driven and can only be achieved by way of schematic assimilation and accommodation. More simply, he characterized humans as "meaning-makers". For psychologists and researchers alike, Piaget's work cast new light on cognition and learning. However, critics such as Vygotsky (1978, 1986, 1987) pointed out the absence of human and cultural mediators in Piaget's theory. Vygotsky argued that symbolic and psychological tools such as linguistics, mathematical systems, and signs all impact the way humans acquire and internalize knowledge. Moreover, parents, teachers, and peers all help to create and alter an individual's schemata through guidance and interaction. Because of this redefinition of environmental factors, constructivism came to be regarded as both an individualized and a collective experience.

The degree to which constructivists have emphasized the interaction between the individual and the environment has led to theoretical diversity. Moving across the continuum, there are those who believe that the learner must never be impeded by external mediators. Instead, the learner should be

allowed to independently explore and discover for him or herself throughout the learning process. It is believed that existing knowledge representations should not be disrupted as they are the property of the individual. Therefore, advocates of this position maintain that information must be accessed, constructed, and encoded in personally meaningful ways in order to be memorable and transferable (e.g., Piaget, 1970; von Glaserfeld, 1984).

Theoreticians who adhere to a more conservative perspective suggest that extensive modeling and guidance from knowledgeable others is imperative. Structured assistance is believed to aid the individual in constructing appropriate understandings and interpretations. While personal knowledge representations have utility, they serve only as the canvas for externally provided truths (e.g., Ausubel, 1968; Bandura, 1986).

The compromise between these two positions can be described as a balance between the individual and all external mediators. That is, while the individual is allowed to discover knowledge, skills, and strategies which are personally meaningful and compatible with prior knowledge representations, learning is never in isolation. Instead, it is suggested that learners be provided with appropriate modeling and scaffolding on an as-needed basis (Pressley et al., 1992). This can be in the form of interpersonal communication or through external prompts or cues. The goal is for learners to gradually internalize knowledge, skills, and strategies which are situationally appropriate and personally meaningful. Because of these factors, this philosophic position has become prevalent in constructivist literature.

According to Moshman (1982), this position may be aptly referred to as "dialectical constructivism". He writes:

For the dialectician, the source of all knowledge lies in the continuing interactions between organism and environment, neither of which can simply impose itself on the other. New knowledge is a constructed synthesis which resolves the inevitable contradictions arising during the course of such interactions. In Piagetian terms, such a synthesis may be seen as an equilibration of assimilatory and accommodatory poles in the dynamic interaction of the inseparable knower and known (p. 375).

Thus, the acquisition of new knowledge and the reorganization of what is already known is motivated by the gap between one's present understanding and the understanding which is required to comprehend the world (Pressley, et al., 1992). By means of oral and written interaction, the learner is obliged to construct his or her own truths by questioning and/or defending present understandings and conceptions both internally and externally (Stein, Bernas, Calicchia, & Wright, 1996). Hence, the term "dialectical" is befitting as it refers to the notion that some ideas will clash or be inconsistent with others (Engestrom, 1987; Paul, 1990). The crucial interplay between knowledge acquisition and knowledge reorganization is guaranteed as the learner is continuously thrust into situations which invoke cognitive conflict and cognitive restructuring.

Situated Cognition and Cognitive Apprenticeship

Based upon the definition of dialectical constructivism given above, comparisons to situated cognition are easily drawn. Rooted in artificial intelligence and cognitive psychology, situated cognition attempts to account for how one learns in a conceptual environment. The conceptual environment consists of the external world as it is perceived, the internal representations of such perceptions, and the resulting interactions (Reynolds, Sinatra, & Jetton, 1996).

As suggested by its name, situated cognition places a premium on situation and context. However, unlike social cognition, greater emphasis is

given to the internal operations which result from an individual's interaction with external mediators (Derry, 1992; Greeno, 1997); that is, the movement which occurs from interpsychological to intrapsychological (Vygotsky, 1986). Still, there are critics who maintain that situated cognition creates knowledge representations which are only acquired in social situations. Furthermore, once created, the representations are domain specific, nontransferable and contextually-bound (e.g., Anderson, Reder, & Simon, 1996). While it is difficult to dismiss these criticisms, they have been touted as exaggerations (Greeno, 1997). A close inspection of cognitive apprenticeship (Collins, Brown & Newman, 1989), an honoree of the philosophies of situated cognition, serves to demonstrate that abstract, transferable knowledge, skills, and strategies can be acquired in social environments which are no more complex than those found within a classroom.

Collins et al. (1989) coined the term "cognitive apprenticeship". Like traditional apprenticeship, an "expert", or more knowledgeable other guides the learner as he or she acquires both complex domain specific as well as abstract processes. This is facilitated by way of modeling, coaching, and scaffolding. Furthermore, the learner is encouraged to articulate and reflect upon his or her own performance as well as the performance of the expert. This incites the learner to consciously access, control, and construct personally meaningful knowledge, skills, and strategies. Finally, exploration is emphasized for the purposes of engendering a state of adaptive motivation and learner autonomy. Thus, the learner is equipped to function effectively within a specific domain as well as to transfer concrete and abstract knowledge, skills, and strategies due to schematic construction and personal exploration.

While cognitive apprenticeship does encourage the importance of

learning within context, it explicitly states that acquired knowledge, skills, and strategies must be decontextualized. Specifically, that it should “extend learning to diverse settings so that students learn how to apply their skills in varied contexts. Moreover, the abstract principles underlying the application of knowledge and skills in different settings should be articulated as fully as possible by the teacher, whenever they arise in different contexts” (Collins et al., 1989, p. 459). Exemplary models of the philosophies of cognitive apprenticeship include Reciprocal Teaching (Brown & Palincsar, 1989; Palincsar & Brown, 1984) and Reciprocal Peer Questioning (King, 1990, 1991a, 1991b, 1992, 1994a, 1994b; King & Rosenshine, 1993).

Reciprocal Teaching

Reciprocal Teaching (Brown & Palincsar, 1989; Palincsar & Brown, 1984) is a step-by-step model which has had considerable success across all learning domains including reading, mathematics, writing instruction, and social studies. The model consists of four critical strategies: questioning, clarifying, summarizing, and predicting. These strategies are acquired and practiced in an environment comprised of cooperative learning, expert scaffolding, and guided instruction. Each strategy has a specific purpose. For instance, question construction leads to a greater integration of text; clarification assists the instructor, the group, and the learner to monitor not only their own comprehension levels, but the comprehension of others; summation promotes further analysis and self-evaluation of the learner’s knowledge, skills, and strategies; and prediction activates prior knowledge (Derry, 1990; Lysynchuk, Pressley, & Vye, 1990). Of particular interest, students were trained to develop questions incorporating “who, what, where, when, why, and how”. These

prompts are referred to as signal words (Rosenshine, Meister, & Chapman, 1996) and have been successful in improving text comprehension.

Empirically, Reciprocal Teaching has been shown to improve reading comprehension scores as well as metacognitive awareness (Palincsar & Brown, 1984; Lysynchuk et al., 1990). Thus, it has succeeded in decontextualizing vital knowledge, skills, and strategies which are required across domains.

Reciprocal Peer Questioning

Like Reciprocal Teaching, Reciprocal Peer Questioning (King, 1989, 1990, 1991a, 1991b, 1992, 1994a, 1994b; King & Rosenshine, 1993) focuses upon the construction of questions and responses and the integration of schematic structures with new knowledge, skills, and strategies.

Reciprocal Peer Questioning begins with explicitly instructing participants on the utility of question generation using generic question stems. Generic question stems require the learner to complete skeletal question outlines. For example, "How are...and...alike?" "What is the main idea of...?" "How does...tie in with what we have learned before?", and so forth (e.g., King, 1989). They require the learner to make connections both within the text and to his or her prior knowledge.

Following question generation, each individual learner is required to independently generate two or three questions relevant to the material being studied. In small, cooperative groups the learners take turns posing their questions to one another. As with Reciprocal Teaching, feedback is provided by the instructor and/or peer group.

According to King (1990, 1991a, 1991b, 1992), this model produces

significantly higher achievement scores than those for discussion alone, questioning and responding without guidance, and independent study. This makes Reciprocal Peer Questioning and generic question stems a promising approach toward the development of cognition, metacognition, and knowledge construction by requiring the learner to activate and utilize prior knowledge, to generate higher-level meaningful questions, and to monitor one's own knowledge, skills, and strategies.

Therefore, in response to the criticisms lodged against situated cognition, it is evident that this theoretical and instructional approach does not invariably lead to the acquisition of contextually-bound knowledge, skills, and strategies. Rather, through the combination of dialectical and constructivist activities, learners can acquire both domain specific and abstract schemata.

Text Comprehension

Studies involving questioning typically employ text. Whether it be narrative or expository, written text has been regarded for centuries as a vital communicative instrument. A relative newcomer in the language tradition, text has proven to be mobile, complex, and enduring by extending one's memories and the number of communicative partners (Goodman, 1984; Olson, 1994). With each piece of text, a writer attempts to convey a message. In the case of narrative text, the theme is typically derived from the author's values, concepts, and life experiences (Goodman, 1984). Expository text, on the other hand, is intended to report or explain information and events. Regardless of the genre, written material is only effectively communicated when its meaning and structure are comprehensible to the reader (e.g., Lorch & van den Broek, 1997; Wandersee, 1988). That is, written text is only effective when it engages one's

needs, interests, and prior knowledge (Goodman, 1984; Mannes & St. George, 1996; Olson, 1994; Tobias, 1994; Wittrock, 1990; Woloshyn et al., 1994).

Over the decades, research has converged on the notion that effective reading comprehension involves the construction of meaning (e.g., Adams, 1990; Bartlett, 1932; Goodman, 1984; Kintsch, 1988; Pressley, Symons, McDaniel, Snyder, & Turnure, 1988; Spires & Donley, 1998; Wittrock, 1990; Wood, Pressley, & Winne, 1990). As early as Bartlett (1932), researchers have asserted that readers construct a meaningful mental representation of text in a twofold manner: a) by conceptualizing the structure of the text and b) by integrating its contents with prior knowledge. Thus, the reader not only seeks a mental representation that is locally and globally coherent, but retrieves and cultivates schemata which serve as the foundation for such constructivist, meaning-producing processes as inference, elaboration, clarification, and prediction (Graesser & Zwaan, 1995; Lorch & van den Broek, 1997; McNamara & Kintsch, 1996; O'Brien, 1995; Wittrock, 1990; Woloshyn et al., 1994; Wood et al., 1990). Without such processing, information is easily forgotten and difficult to process.

To more clearly understand constructivism and reading comprehension, it is important to understand connectionist theory. Connectionist theorists depict knowledge, skills, and strategies as belonging to associative networks or schemata (McClelland & Rumelhart, 1986; Rumelhart, Hinton, & McClelland, 1986). Developed by way of simultaneous processing, schemata are depicted as abstract data structures, patterns, prototypes, and internal models which encompass the relationships between concepts, ideas, procedures, and so forth. McClelland and Rumelhart (1986) state that "...the processing system is

assumed to consist of a highly interconnected network of units that take on activation values and communicate with other units by sending signals modulated by weights associated with the connections between the units...." (p. 173). More simply, units, which may number anywhere from thousands to millions, are organized into modules, or schemata. When activated, the units within each module interact internally. This process generates and/or strengthens their connectivity. These same units also communicate externally with units inherent within other modules. Consequently, each module is a synthesis of the states of all of the modules from which it receives inputs.

The marriage of constructivism and connectionism within reading comprehension has been supported by several comprehension theorists (Gernsbacher, 1996; Graesser, Swamer, Baggett, & Sell, 1996; Mannes & St. George, 1996; Morris, 1990; Sharkey, 1990). Perhaps the most renowned proponent is Kintsch (1988) and his construction-integration (CI) model of comprehension which is based upon a connectionist algorithm. Kintsch (1988, 1992) suggests that processing occurs in two stages. During the first stage (i.e. construction), concepts from the text, as well as syntax, semantic, and world knowledge are activated to produce a network, or module of activated concepts. The activation process then continues throughout the network strengthening links between contextually compatible units or concepts and weakening those which are incompatible (i.e. integration). This continues until a stable state is achieved. The outcome is referred to as the reader's mental representation (Kintsch, 1988, 1994a, 1994b; McNamara & Kintsch, 1996).

Although it is a unitary structure, a mental representation embodies a number of different components. Of primary importance are the textbase and situational models. Textbase construction involves the extraction of semantic

information from a text (McNamara et al., 1996). It also involves the propositionalization of text contents so that a reader may develop a sense of text coherence at both local and global levels. At the local level this is accomplished by carrying propositions over from one processing unit or cycle to the next, or from one sentence to the next. At the global level, coherence is enhanced by the appropriate reinstatement in the textbase of propositions from a preceding part of the text. The result being a locally and globally well-structured memory representation of the text. This textbase representation will enable readers to verify statements they have read, answer questions about the text, and to recall and summarize the text (Kintsch, 1994a; McNamara et al., 1996). Consequently, to develop a coherent, memorable representation of the textbase, both the writer and the reader must activate the most important propositions so that they are frequently processed and recently stored in memory.

While the textbase is the mental representation of the text itself, the situation model is the mental representation of the situation described within the text (Kintsch, 1989; 1994c). It is constructed using the textbase as well as the reader's prior knowledge and experience. Various sources contribute to the situation model: knowledge about the language of the text, world knowledge, and knowledge concerning the specific communicative situation (McNamara & Kintsch, 1996). According to Kintsch (1988), what separates text memory from text learning is the activation of prior knowledge and the development of a situation model (Schmalhofer, 1996). More simply, learning involves the activation of knowledge networks which are subsequently elaborated upon using the information embodied within the text (Kintsch, 1994a; McNamara & Kintsch, 1996). Text memory, on the other hand, involves the development of a

coherent, well-structured textbase model which may, at best, be loosely linked to the reader's prior knowledge. Information processing theorists such as Sternberg (1985, 1988) and Mayer (1989, 1992) share a similar learning and memory viewpoint. Like the textbase model, selective combination combines "selectively encoded information in such a way as to form an integrated...internally connected whole" (Sternberg, 1985, p. 107). Selective comparison bears similarity to Kintsch's situation model by relating "newly acquired or retrieved information...to old knowledge so as to form an externally connected whole" (Sternberg, 1985, p. 107). Mayer (1989, 1992) refers to these processes as internal connections and external connections, respectively.

Kintsch (1994a) also included text structure as an integral component of text comprehension. The first type he calls microstructure. It is comprised of local text properties, or explicit indicators of relations between concepts and ideas (i.e., connectives, argument overlap, pronominal reference), which give the text some structure at a propositional level (McNamara & Kintsch, 1996). A strong structure is imperative as the comprehension of propositions is the cornerstone to discourse processing (Murdock, 1995). Conversely, macrostructure refers to the global organization of the text. Derived from the microstructure through mapping rules (e.g., deletion, generalization, and construction), macrostructure is the hierarchically ordered set of propositions which represent the global structure of the text. At the helm of the hierarchy are macropropositions (i.e., the main points) and their hierarchical relations. Macrostructure is sometimes explicitly signaled by organization signals (e.g., headings, topical overviews, topical summaries) (Lorch & Lorch, 1996), but is most often inferred by the reader (Kintsch, 1996). Inference is achieved through a reduction process characterized as an abstraction or summarization

of text (van Dijk, 1995).

According to the literature, the manner in which microstructure (e.g., Kintsch & van Dijk, 1978; McKoon & Ratcliff, 1992, 1995; Murdock, 1995) and macrostructure are constructed (e.g., Albrecht et al., 1995; Albrecht & O'Brien, 1993; Beck & McKeown, 1991; Beck, McKeown, Sinatra et al., 1991; E. Kintsch, 1990; Kintsch & van Dijk, 1978) are vital components of text comprehension and learning. The better organized they are, the more a reader understands and remembers the text. This is especially true of macrostructure as it inculcates a global understanding of text meaning. Moreover, it helps to define microstructure since the establishment of local coherence involves macrocontrol in the form of a theme, topic, or point (van Dijk & Kintsch, 1983). That is, good microstructure requires discourse referents to be ordered relative to a central referent (i.e., a person or object), predicate sets to be organized according to major predicates (i.e., macroactions or macroevents), properties of individuals to amount to global property, and so forth. Similarly, sequences of actions or events necessitate global goals and motivations. And finally, there must be a unity of time and place for these events, actions, and participants. Simply put, pronominal inferences, bridging inferences, causal connections, and argument overlap are not only vital to local coherence, but are inextricably linked to global coherence. Thus, microstructure and macrostructure share an integral relationship.

As suggested, text comprehension involves microprocessing and macroprocessing. According to most theorists (e.g., McKoon & Ratcliff, 1992; van den Broek et al., 1995; van Dijk & Kintsch, 1983), the reader routinely seeks to establish local coherence. This is accomplished through the linkage of noun and pronoun arguments in the current clause with explicit arguments and

propositions in the previous text (Graesser, Bertus, & Magliano, 1995; McKoon & Ratcliff, 1992; O'Brien, 1995); attending to connectives (i.e., and, or, because, so) which link adjacent clauses and clauses co-occurring in working memory; and inferring causal chains and relations between clauses that co-occur in the working memory (Graesser, Bertus, & Magliano, 1995). In regards to macroprocessing, whether text is poorly structured or well-structured, readers must engage the processes of deletion, generalization, and construction in order to infer macrostructure. In other words, readers must select those propositions necessary to garner the gist or upshot of the text, compare the selected propositions against subsequent propositions, and construct the remaining important propositions into a text schema. Having engaged in this process, text information will be efficiently organized within memory and, depending upon the knowledge base and skill of the reader, integrated with previous schematic structures. The proficiency of this process and the quality of the end-product will play an important role in memory retrieval and knowledge transfer.

While macrostructure is occasionally signaled within the text (Kintsch, 1998; Beck & McKeown, 1991; Beck, McKeown, & Gromoll, 1989; Beck, McKeown, Sinatra et al., 1991) it is typically the reader's responsibility to infer its existence. This is a rather formidable task for poor comprehenders, young readers, and those with low domain knowledge. The same is true for microstructure as poorly constructed text at the microlevel can impede comprehension for those with weak comprehension skills (i.e., poor word recognition, working memory deficits, and delays in perceptual speed and automaticity) (e.g., Adams, 1994) and deficient knowledge. Due to these potential deficits, expectations of unassisted local coherence and

macrostructure construction may be naive. What may be required is structured scaffolding in the form of prompts or cues.

With respect to younger readers, their macroprocessing skills are only just emerging. Studies have shown that there is a tendency among school-age students to process text in a linear, element-by-element fashion (e.g., E. Kintsch, 1990; Scardamalia & Bereiter, 1984). This is especially true in regards to expository text comprehension as it is often unfamiliar to the reader in terms of content and structure (e.g., Beck & McKeown, 1991). Any unfamiliarity can encumber readers from abstracting the general meaning or gist of the text, leaving them to process the material in a sentence-by-sentence manner (E. Kintsch, 1990). Such linear processing typically results in shallower representations (Bereiter & Scardamalia, 1987; Scardamalia & Bereiter, 1984) compared to the multilayered macrostructure representations derived by older, more proficient readers (E. Kintsch, 1990). This can be exacerbated by the surprising number of poorly structured instructional texts (e.g., Beck & McKeown, 1989; Beck, McKeown, Sinatra et al., 1991; Britton & Gulgoz, 1991).

Beck and McKeown (1989) examined the expositions found in Grades 3 through 6 of recent editions of four basal Readers. What they found were local and global coherency breaks. Their use of coherence as a construct involved two broad categories of relationships: a) the relationship of the part to the whole (i.e., how a specific section relates to the main topic of the selection) and b) the relationships among the parts (i.e., sentence by sentence, paragraph by paragraph, the selection provides an overall organization). What they found were frequent disruptions by the introductory passage and by internal passages with digressions to a new topic; an excessive number of subtopics; loose collections of ideas that revolved around a single topic; an absence of

overarching concepts; poorly categorized materials; headings which misrepresented the nature of the material; and a tendency to disrupt the flow of ideas. Similar findings arose during Beck, McKeown, Sinatra, and Loxterman's (1991) review of a fifth grade social studies text and Beck, McKeown, & Gromoll's (1989) examination of four widely used elementary school social studies programs and their accompanying textbooks. Therefore, the assumption that readers have sufficient background knowledge which can override inadequate explanations, ill-defined relationships, and poor organization appears to be widespread. In response, Beck and colleagues (1989; 1991) state that expository text must enable students to build mental models by clearly presenting the goals of the content. That is, authors need to portray the situation and textbase of the written work in an explicit fashion. Moreover, students need opportunities to clarify and elaborate their conceptions of text content. While this is certainly true for most readers, it is especially true for those lacking sufficient background knowledge and comprehension skills.

In an attempt to demonstrate how comprehension can be improved by changing these and other coherency breaks, Beck, McKeown, Sinatra, and Loxterman (1991) revised four segments of Grade 4 and 5 American history textbook. Revisions primarily involved repairs to causation/explanation. That is, the text was inconsiderate of the connections between causes to events and events to consequences. The general operations used in making the revisions included clarifying, elaborating, explaining, and providing motivation for important information. Moreover, connections were made more explicit. The results of the study indicated that those students who read the revised text recalled more information and answered more questions correctly than their original text counterparts. There were also differences in understanding,

garnered by way of recall protocol analysis and question responses related to specific ideas in the text, suggesting that students in the revised text group had a superior understanding of text content.

Micro- and macrolevel text revisions can be especially helpful for younger students with low domain knowledge. McNamara et al. (1996) conducted two experiments concerning the role of text coherence in the comprehension of science texts. Subjects in both experiments were junior high school students. Experiment 1 sought to compare the individual contributions of local, global, and explanatory coherence of biology text on textbase and situational understandings. While local and global coherence have been previously explained in this discussion, explanatory coherence refers to the "content that supplies background knowledge needed to understand the text that the reader may not have" (p. 6). Therefore, three versions of a chapter regarding mammalian traits were employed: a) an original version which was locally coherent and globally incoherent; b) a revised version which was made globally coherent with explicit macrosignals; and c) an expanded version which added explanatory coherence to the text. Measures included a post-test comprised of multiple choice, true-false, fill-in-the-blank, and short answer questions (indicative of both textbase and situation model development), text recall (indicative of textbase model development), and a sorting task (indicative of situation model development). Results indicated that students who had read the revised version recalled more of the text than those who read the original version. Changes in sorting patterns were also largest for those who had read the revised version of the text.

These findings led the researchers to question whether domain knowledge had an impact on performance scores. A number of students had

superior scores using the original text. Such a finding caused the researchers to question whether this was due to high domain knowledge. Thus, Experiment 2 sought to test the prediction that some students may learn better when they are forced to make the text more coherent at local and global levels. Four versions of a text regarding heart disease were subsequently created: a) maximum coherence at the local and global levels (CM); b) maximum coherence at the local level and minimum coherence at the global level (Cm); c) minimum coherence at the local level and maximum coherence at the global level (cM); and d) minimum coherence at both the local and global levels (cm). The same testing procedures as in Experiment 1 were used.

Text recall results indicated that high knowledge readers performed consistently better than their low knowledge counterparts. Of those with high knowledge, coherence at either the global or local levels made little difference. Subjects were able to construct a good textbase in spite of what was present or absent in the text. For low knowledge subjects, coherence was required at either the global or local level for textbase construction. The poorest performance was for low knowledge readers using the cm text.

The post-test question indicated that high knowledge subjects performed better after reading the cm and CM texts than the other two reading conditions, and better on the cm overall. The questions they performed best on were problem solving and bridging inference. This was the reverse for low knowledge subjects as they not only performed most poorly on problem solving and bridging inference questions, they also had the worst overall performance when given the cm text. In regards to textbase questions, all participants who read the globally coherent text answered these questions more accurately than those who read the texts with poor global coherence. For changes in sorting

scores, cm text was most effective in changing sorting patterns of high knowledge subjects while those texts with good local coherence (Cm and CM) were most effective for low knowledge subjects.

It therefore appears that the more active processing required by high knowledge subjects, the better their performance scores. Any lack of explicitness incited them to construct not only textbase models, but strong situation models. This was aptly demonstrated on the bridging inference and problem solving questions of the post test and in the sorting task. This pattern was reversed for low knowledge subjects. The results for this group of students indicated that highly explicit text was required to comprehend and retain information from the text. Without local and/or global coherence, low knowledge readers were unable to fill-in the missing information necessary to construct either a textbase model or a situation model.

Similar results were found by McNamara and Kintsch (1996) in a study involving college-aged students. Only two versions of text were employed, high coherence (CM) and low coherence (cm). The results indicated that low knowledge readers studying low coherence expository text engaged in longer reading times and responded more poorly to open-ended questions when compared to low knowledge readers studying high coherence text, high knowledge readers studying low coherence text, and high knowledge readers studying high coherence text. High knowledge readers studying low coherence text generated the most superior performance results on both sorting tasks and open-ended question response. The conclusion drawn by the researchers is that low coherence text must induce high knowledge readers to engage in more intense inferential processing. The high knowledge reader is incited to process the material at a deeper, more meaningful level (Craik & Lockhart, 1972).

To better understand these results, if text coherency is both locally and globally poor, comprehension may be impeded by hindering the development of an effective textbase model. A strong textbase model is typically required to build a strong situation model. If, however, the reader brings to the text ample prior knowledge, it is possible that a strong situation model may arise which has the potential to override comprehensibility shortcomings. That is, the reader has sufficient prior knowledge to fill-in structural gaps and to construct a coherent mental representation of the text topic or theme. A strong situation model is also attainable for the high knowledge reader if either local or global coherence alone is poor. This is not the case for the low knowledge reader. For instance, if text is locally incoherent but globally coherent, a low knowledge reader may be able to create a textbase model, but it is unlikely to be very strong since the information necessary to fill-in structural gaps is unavailable. In regards to situation model development, even the strongest global structure would not be enough given the reader's lack of domain knowledge. If the text is locally coherent but globally incoherent, situation model development would still be difficult. As indicated earlier, a strong macrostructure (i.e., global coherence) is required to comprehend the upshot or gist of the text. When coupled with low domain knowledge, a weak understanding of the text topic would make situation model construction problematic. The reader would, however, have the opportunity to infer a strong textbase model because the text is locally coherent. Therefore, according to the construction-integration model, text comprehension is not only dependent upon the knowledge of the reader, it is additionally influenced by text structure and coherency.

In summary, the conclusions drawn by McNamara et al. (1996) are that reading should be challenging enough to the reader to stimulate active

processing, but not so challenging as to confuse the reader. The researchers also suggest that textbook "customizing" may be a suitable option. That is, on the basis of a subject's knowledge base, instructional text could be manipulated at the level of coherence to ensure reader inferencing. While this suggestion appears logical, it also seems impractical as it would be extremely costly and ill-suited to regular classroom activities. After all, far too many books in classrooms and libraries are lacking local and/or global coherency. To discard all of these books would be impossible. Moreover, as demonstrated, each student has a varying degree of previous knowledge. Accurately matching one's level of prior knowledge to a textbook would be labour intensive, expensive, and unreliable. A more feasible solution is to find a reading/learning strategy which induces active or deep processing, regardless of the level of one's prior knowledge or text structure. One such approach may be questioning as there is evidence to suggest that posing and answering questions helps to build internal (textbase) and external (situational) connections (Mayer, 1989, 1992). In adopting a specific questioning technique, the reader is given a scaffold by which a coherent textbase model can be constructed and subsequently linked to as much prior knowledge as possible.

Therefore, it behooves educational researchers to not only suggest methods which are universal to any reading scenario/environment or level of domain knowledge, but that are cost efficient, transportable, and controllable by the reader.

Measurement of Text Comprehension and Knowledge Acquisition

The most common purposes for text comprehension measurement are to analyze either comprehension processes (i.e., word recognition,

comprehension monitoring, text scanning strategies) (e.g., Adams, 1990; Sattler, 1992) or comprehension products or outcomes. Much of the latter shares theoretical and practical philosophies with the discipline of knowledge assessment. The means by which knowledge is acquired, structured, and retrieved is inextricably linked to all learning and memory domains, including text comprehension. Thus, many of the same assessment methods are employed.

An informal survey of recent text processing and knowledge acquisition literature suggests that the most common methods of assessment are recall (Britton & Gulgoz, 1991; Britton, Van Dusen, Gulgoz, & Glynn, 1989; Cote, Goldman, & Saul, 1998; Estes, 1995; Lawson & Chinnappan, 1994; Lorch, Lorch, & Inman, 1993; Machiels-Bongaerts, Schmidt, & Boshuizen, 1995; Mannes & Kintsch, 1987; Mannes & St. George, 1996; McKeown, Beck, Sinatra, & Loxterman, 1992; Moravcsik & Kintsch, 1993; Schmidt, De Volder, De Grave, Moust, & Patel, 1989;) and multiple choice (Britton & Gulgoz, 1991; Britton, Van Dusen, Gulgoz, & Glynn, 1989). In many cases these forms of assessment are used in combination with each another or with other forms of assessment. However, they are also employed in isolation leaving some scholars concerned for their validity.

One of the criticisms lodged against recall of declarative or procedural knowledge and other highly structured formats such as multiple choice is that they provide little indication either of the level at which students understand the subject matter or the quality of their thinking (e.g., Nickerson, 1989; Norris, 1989). This is particularly true of single-right-answer formats (i.e., multiple choice). The purpose of an instrument should not be simplified scoring procedures, but the measurement of the nature and power of the student's

organized structure of knowledge (Goldsmith, Johnson, & Acton, 1991). After all, one's knowledge structure either facilitates or hinders what one can do within a subject area (Messick, 1984). Such structures encompass broader cognitive skills or abilities that are applied to the subject matter enabling the individual to remember, visualize, interpret, transform, evaluate and think both convergently and divergently about the material (de Jong & Ferguson-Hessler, 1996; Messick, 1984). Being able to access these skills and abilities poses great difficulties for researchers. According to some theorists, such complex knowledge structures should be subject to direct as well as indirect assessment (Kluwe, 1993; Kintsch, 1998; Messick, 1984; Putz-Osterloh, 1993). Because both forms of assessment have their limitations, it is vital to obtain results from diverse methods for the purposes of confirmation (Kintsch, 1998).

In the case of direct assessment methods, they do not allow the researcher to access all that the subject may know or understand. Furthermore, some students have developed strategies which allow them to generate acceptable answers without a deep understanding or to answer questions with little understanding. Finally, due to a dearth in question answering theories, a reliable means of constructing the right questions is lacking (Goldsmith et al., 1991; Kintsch, 1998). Therefore, while direct assessment approaches such as short answers and summary and/or recall questions have their merits (i.e., accessing the phenomenology of the individual, serving as an appropriate means of text memory) and are considered functional, they should not be considered conclusive. In response to these criticisms, many text comprehension and knowledge assessment researchers are employing indirect methods, either exclusively or in combination with direct approaches. The most common trend is toward scaling methods.

Scaling methods contain a set of key words or phrases which are characteristic of a certain knowledge domain. Subjects are required to organize or judge the words in a manner befitting the task. The organizations are then compared to those constructed by experts. Thus, the purpose of the procedure is to reveal the knowledge structure of each subject without conscious interference.

The most basic scaling technique is to instruct subjects to make relatedness judgments between pairs of key words in a set. An example of this technique was employed by Britton and Gulgoz (1991). They selected 12 important terms from a piece of text and constructed all possible pairs using these words. This resulted in 66 pairs in total. Each subject was then asked to rate each pair for relatedness using a 7-point scale. These ratings were perceived as quantitative representations for each subjects' ideas about the relationships between the 12 terms. The representations were then compared to the intended representations: those of the author and the 7 subject-matter experts. If the correlation correspondence was high, the subjects were presumed to have a strong mental representation. The relatedness ratings were additionally interpreted as distance measures which are then used to create a map of the structures underlying them. A number of choices were available to the researchers (i.e., multidimensional scaling, cluster analysis, tree analysis, and network analysis). Because the 12 terms appeared to be arranged in a network fashion as opposed to dimensions, clusters, or trees, a network analysis was conducted.

While relatedness judgments have some utility, they are also limited in their usefulness. Pairwise judgments can be too laborious for subjects, particularly when the number of keywords increases. Thus, they are impractical

if other assessment measures are to be used or if time is a consideration. Moreover, the majority of knowledge domains are too complex to be restricted within the space of simple pairwise dimensions (Kintsch, 1998). In response to these concerns, Kintsch (1998) and Ferstl and Kintsch (1999) suggest two indirect knowledge assessment techniques. They maintain that when measuring the amount of learning that occurs from reading a text, the researcher must ascertain whether there has been an effect on the reader's memory and knowledge. That is, the effect will be whether the reader organized his or her knowledge in the direction of text organization. To determine whether this occurred, the reader's knowledge regarding a particular domain must be assessed. Following this, the reader is directed to study a related text and is then reassessed to see whether his or her knowledge organization has changed in accordance with the organization of the text. The two techniques which conform to this method are cued association and the sorting task.

Cued association requires subjects to silently read a list of key words several times. They are then instructed to generate up to 3 associations for each keyword. No restrictions are placed on the associations. The number of times a keyword is given as a response to another keyword is used as measure of relatedness between the words. The matrix resulting from this procedure is asymmetric which warrants the Pathfinder analysis since it is particularly suited to such data. Conversely, the matrix generated from sorting task data is symmetric. This task requires subjects to sort through cards containing key words and to place those key words into piles according to relatedness. Subjects are told that there is no correct or incorrect way to sort. Instead, the purpose is to better understand the intuitions of the subjects regarding word

relations. Subject matrices are then individually compared with an expert-generated matrix to produce a sorting score for each subject. This scoring method is based on principles of the harmony theory proposed by Smolensky (1986) (as cited in McNamara & Kintsch, 1996). Scores can then be analyzed by way of hierarchical clustering analysis (Johnson, 1967) to determine group trends.

While the sorting procedure and cluster analysis yield readily interpretable results, Kintsch (1998) argues that the results may be too neat suggesting that subjects may have engaged in semantic analysis when sorting. Instead, he proposes that cued association data may be more accurate in reflecting knowledge organization than the less spontaneous sorting. However, a closer look reveals that unlike the sorting task, cued association data is not analyzed to determine a harmony score. Thus, it is simply the subject's perception of how the words are associated as compared to a subject's perception in accordance with that of an expert. Such comparative information can be useful as it verifies whether a subject's knowledge structure resembles that of the author's and/or an expert's. Other merits of the sorting task include its time efficiency, ease of scoring, and ease of interpretability. These factors ensure that the sorting task is a useful, indirect assessment technique which can be combined with other forms of measurement. McNamara and Kintsch (1996) suggest that the sorting task is an ideal measure when combined with other comprehension measures that compete for valuable subject time. Thus, as indicated earlier, employing a combination of direct and indirect methods is imperative when measuring knowledge acquisition and comprehension.

Also imperative is the fit between the knowledge being measured and the knowledge construct. That is, all measurement instruments must be

compatible with the construct or theoretical model being studied. In the case of the construction-integration model (CI) (e.g., Kintsch, 1988), by way of comprehension construction and integration processes, a reader achieves a stable, mental representation of the text being read. This mental representation contains a number of different components. Of primary importance are the textbase and situation models. Therefore, in order to measure a reader's mental representation it is necessary to access his or her textbase and situation models. This can be accomplished by determining the sub-components of each model, developing accompanying questions, and measuring the reader's response(s) to them. For instance, a textbase model is comprised of localized, textbase understandings while the situation model is comprised of bridging inferences, elaborative inferences, and knowledge transferability. To access these sub-components, a direct, open-ended method of assessment may be most appropriate. To indirectly assess a situation model, either the sorting task, relatedness task, or cued association task may be most appropriate.

When studying the effects of text structure on knowledge acquisition and comprehension, it is important to examine its two main components: microstructure and macrostructure. As stated earlier, microstructure is comprised of local text properties that provide structure at a propositional level. Conversely, macrostructure refers to the global organization of the text. As such, it is comprised of macropropositions and their hierarchical relations as well as occasionally provided textual signals. In regards to assessing a reader's conception of microstructure and macrostructure, a summary and/or recall task appears most suitable. This task instructs the reader to summarize and/or recall both the contents and structure of the text. The goal is to provide insight into the number and type of propositions/macropropositions recalled by

the reader and the manner in which they are organized.

Taken together, it is evident that the measurement of knowledge acquisition and text comprehension are complicated ventures. Consideration must be given to several factors prior to selecting the most suitable instruments. They include time constraints, the target population, and most importantly, theoretical underpinnings and related constructs.

Expository Text versus Narrative Text

Situation model construction is of particular importance during expository text comprehension. It is typically assumed that when reading expository text, learners are unfamiliar either with the domain of study or with certain aspects of the domain (e.g. Britton, Glynn, & Smith, 1985). This can in turn lead to comprehension difficulties. Narrative text is rarely undermined by this phenomenon (Lorch & van den Broek, 1997). For instance, it has been shown that the development of narrative schemata in children precedes the development of expository schemata (Voss & Bisanz, 1985). Narrative schemata are typically based upon situational world knowledge (i.e., knowledge about human actions, physical events, and human reasoning) (Black, 1985). Thus, inferences regarding characters, activities, and events are easily generated (Graesser & Kreuz, 1993; van Dijk & Kintsch, 1983). Even in the case of poorly constructed story schemata, subjects typically recall story events in a prototypical order despite sequencing or the recall instructions (Voss & Bisanz, 1985). Based on these results, it appears that when reading narrative text, schema activation and acquisition are unlikely to be problematic.

This is not always the case when reading expository text (Britton et al., 1985; Kintsch, 1988; Lorch & Lorch, 1996). Sufficient background knowledge is

often lacking when entering an expository text situation, which is then exacerbated by typical expository text structure. It has been noted that transitions between expository text topics often fail to involve strong connections between successive topics (Lorch, 1995). Writers frequently assume that readers can fill-in missing information and establish coherence (Graesser, Bertus, & Magliano, 1995). For example, when discussing a specific animal, expository text structure tends to move from describing the animal's place of origin to its eating habits. If the reader does not note these transitions and encode the text topics and their organization, recall is likely to suffer as the reader is unable to access some or all of the topics in memory (Lorch & Lorch, 1996; van den Broek, Risdien, Fletcher, & Thurlow, 1996). The reader is then required to concurrently carry out a variety of component reading processes and memory management processes (Britton et al., 1985) for comprehension to occur. Component reading processes include word recognition, the retrieval of word meanings, sentence parsing, and text integration.

As previously mentioned, text integration is the most important process as it generates the cognitive structures which are the "desired end products of reading" (Britton et al., p. 227). However, due to the limited capacities of the short-term memory, it cannot simultaneously hold all of the component processes and prior knowledge necessary for effective comprehension. What are then required are memory management processes. They involve the rapid shifting of cognitive programs and prior knowledge in and out of the working memory as required by the various processes being executed (Britton et al., 1985). What activates these vital comprehension processes are text features such as headings, topical overviews, and topical summaries. These components of text macrostructure guide the reader to effectively represent and

retrieve text (Lorch & Lorch, 1996).

Despite the finding that ample prior knowledge is an important factor in facilitating text processing (Kintsch, 1988, 1994a), we cannot conclude that poor background knowledge is deleterious. Extending the scenario cited in a previous section, coherent, well-structured text can improve reading comprehension outcomes, in spite of an insufficient knowledge base, due to its explicitness and incitement of general comprehension strategies (Kintsch, 1994a; McNamara & Kintsch, 1996; McNamara et al., 1996; Moravcsik & Kintsch, 1993). To understand this statement it is necessary to explain text coherence.

Coherence at the local level is established when propositions are carried over from one processing unit or cycle to the next, or from one sentence to the next. This process generally occurs between one and three sentences (e.g., Graesser, Bertus, & Magliano, 1995; Kintsch, 1994a; O'Brien, 1995). Links are also made between nouns and pronoun arguments in the current clause with explicit arguments and propositions in the previous text. Furthermore, connectives (i.e., *and*, *or*, *because*, *so*) link adjacent clauses and clauses which co-occur in the working memory (Graesser et al., 1995). Causal chains and connections, bridging inferences and argument repetition (McKoon & Ratcliff, 1992; van den Broek et al., 1995) are also vital components which carefully constructed by the writer and are inferred by the reader when establishing local coherence. Therefore, when text is locally coherent, it is explicit and easier to comprehend.

Coherence at the global level is enhanced by the appropriate reinstatement in the textbase of propositions from a preceding part of the text. This coherency building process involves the organization of local chunks of

information into higher order chunks and the linking of an incoming clause to a clause which occurred much earlier in the text but is no longer available in active memory (e.g., Graesser et al., 1995; Kintsch, 1994a; O'Brien, 1995; Voss & Bisanz, 1985). Like its local counterpart, global coherence requires macrolevel bridging inferences, causal connections, and in particular, argument repetition (Tapiero & Denhiere, 1996). Finally, structural contributors to global coherence involve macrostructure signals such as headings, topical overviews, and topical summaries, as mentioned above. Thus, globally coherent text has overall unity and the theme or topic is apparent to the reader.

Therefore, despite an insufficient prior knowledge-base, as may be the case when reading expository text, it is evident that information can be effectually processed and learned if it is well-structured and coherent. Support for this position can be found in research conducted on both adult (Britton & Gulgoz, 1991; McNamara & Kintsch, 1996) and adolescent readers (McNamara et al., 1996).

Expository Text and Questioning

To date, there has been little research concerning expository text comprehension and questioning techniques. Typically, what has been conducted has involved either narrative text (e.g., Graesser, Robertson, & Anderson, 1981; Graesser, Robertson, Lovelace, & Swinehart, 1980) or lecture/classroom lessons (King, 1989, 1990a, 1991b, 1992a, 1992b, 1994a, 1994b; King & Rosenshine, 1993). Nevertheless, there are a few notable exceptions. Davey and McBride (1986), Andre and Anderson (1978-1979) and Lysynchuk et al. (1990) have researched expository text and questioning techniques with school-aged subjects. Only Davey and McBride (1986)

considered prior knowledge when designing their research. They selected four passages which were in accord with fifth-grade student interest levels (moderately high), decoding capabilities (individual words were pronounceable with 98% accuracy for the subjects' reading levels), and topic familiarity (moderate to low). Similar to the other studies cited, they employed well-structured text. To this author's knowledge, there are no published research findings which address the combination of questioning techniques, low prior knowledge and ill-structured expository text. The utility of such research is evident as not all text is well-structured and few learners approach new learning tasks laden with sufficient prior knowledge.

Davey and McBride (1986) studied the effects of training in question generation on comprehension question performance, on the quality and form of generated questions, and on the accuracy of predicted comprehension. Using a stratified random procedure based upon two reading achievement scores (California Achievement Test, 1977, Reading Comprehension subtest), 125 sixth-grade subjects were assigned, to five experimental groups: question training (QT), no-question control (NQC), question-generation practice (GP), inference question practice (IP), and literal question practice (LP). Materials consisted of 15 training and practice passages and 4 test passages. Each test passage was accompanied by 8 free response comprehension questions. The passages were between 240-280 words in length and were written at a fifth-grade reading level. As mentioned, they were expository passages and had been pilot tested with subjects who were similar in both age and skill level.

The training given to each condition varied significantly. The QT group had been trained to generate two types of questions, those linking information across sentences and those tapping the most important information. The IP and

LP groups practiced responding only to inferential questions and literal questions, respectively. They were not trained to generate questions, only to practice answering specific types of questions. The questions were based upon the provided passages. The GP group was instructed to generate two good think-type questions for the passages. The guidelines for generating questions were straightforward; that is, that good questions assessed the most important ideas in a passage, that they should make the reader think about what he or she had read, and that they could not be answered by merely underlining parts of a passage. Finally, the NQC was instructed to read the same passages as the other conditions but to complete a vocabulary activity instead of generating or answering questions. The activity involved thinking about the meaning of underlined words from each passage and finding their definitions.

Subjects were assessed across two testing sessions. They were instructed to read two passages per sessions and to generate two think-type questions comprised of the most important information in each passage. Similar to the instructions provided for the GP, the subjects were informed that good questions made the reader think about what they were reading and that the answers could not be underlined in the passage. Following this, the subjects were to respond to their 4 inferential and 4 literal questions without looking back to the text. Once this was completed they were to rate how well they believed they had done on the questions using a 4-point Likert scale. There were no time limits.

After 5 experimental sessions, the QT group outperformed the 4 comparisons conditions on measures determining the quality of the generated questions; all but the IP group on the appropriateness of generated question form; the NQC and IP on literal test passage comprehension and all other

groups on inferential test passage comprehension; and all groups on the accuracy of predicted response performance on the comprehension questions. Based upon these results, it is evident that question training improves the comprehension of expository text, even when prior knowledge is moderate to low. This is because both the instructions and training given to the QT subjects enabled them to bridge existing gaps within the text as well as to generate vital internal connections.

What makes this study so intriguing is that the questions were generated by the subjects and not the experimenters. While it may be argued that one's response to his or her own question can be expected to be correct, the questions had already been scored for quality by the experimenters. If the question required a text-derived inferencing of the macrostructure type (e.g., central ideas or gist) or of the text-connecting type (e.g., integrating information across sentences) it was graded as correct. However, if the question required a response which was a direct restatement of text information, or if it required evaluation based upon the reader's attitude, prior knowledge, or a combination of both and not an application of passage information, it was scored as incorrect. Interrater reliability was .89 based upon data collected from two examiners. Therefore, certain kinds of self-generated questions can produce responses which demonstrate a clear understanding of text content.

However, as discussed by Kintsch (1994a), there are fundamental differences between memory and learning. It is his assertion that remembering a text and learning from it are separate issues. "Remembering a text means that one can reproduce it in some form, more or less verbatim and more or less completely, at least its gist. Learning from a text implies that one is able to use the information provided by the text in other ways, not just for reproduction." He

goes on to add that "learning requires deep understanding of the subject matter, so that the information acquired can be used productively in novel environments; for mere memory, as assessed by reproduction of the text, a more shallow understanding suffices" (p. 294). Therefore, while it is evident that the QT subjects within the Davey and McBride study remembered the text, according to Kintsch's definition we cannot assume that they learned the material as their questions and responses were to link information across sentences and tap the most important ideas. Consequently, the knowledge generated by the QT group was not tested for application or transfer effects. Instead, the self-generated questions and responses in the QT condition were to demonstrate superior reading comprehension skills as compared to those generated by the other experimental conditions. More research is required to ascertain whether self-generated questioning leads to superior learning as opposed to text memory.

While Davey and McBride (1986) were able to demonstrate that self-questioning can lead to the successful reading comprehension of well-structured expository text, even when prior knowledge is low to moderate, Andre and Anderson (1978-1979) and Lysynchuk et al. (1990) demonstrated that self-questioning of well-structured text leads to successful comprehension without concern for the existence of prior knowledge. That is, there were no attempts to control for prior knowledge when selecting expository passages in either study. This makes the assumption that the text material was remembered and/or learned difficult to support (Wong, 1985) as it may simply be an integration or elaboration of previously known information (Mandler, 1985, 1989). Nevertheless, the studies were able to demonstrate that with training, self-questioning leads to superior reading comprehension scores when

compared to students with *read-reread* strategy training (Andre et al., 1978-1979) or no training at all (Lysynchuk et al., 1990).

It is important to note that the results of Lysynchuk and colleagues must be viewed with caution since their approach involved not only self-questioning, but also prediction, clarification, and summarization; the components of reciprocal teaching (Palincsar & Brown, 1984; Brown & Palincsar, 1989). Because of this, it is difficult to verify whether self-questioning was the primary variable facilitating successful reading comprehension. Nonetheless, as mentioned, there is sufficient evidence to demonstrate the salience of self-questioning in areas other than text processing (i.e., lesson and lecture comprehension).

Generic Question Stems

Alison King has successfully shown that self-questioning improves both lesson and lecture comprehension, even in the absence of prior knowledge (1989, 1991b, 1992b, 1993, 1994). Embedded within an approach entitled Reciprocal Peer Questioning, subjects are trained to generate questions based upon lecture and lesson content using generic questions stems. Upon completion of this task, they are then to assemble in small groups to collectively pose and answer one another's self-generated questions. It should be noted that while verbalization and group interaction may add to the effects of self-questioning, previous concerns regarding Reciprocal Teaching do not pertain to this discussion since, unlike its counterpart, Reciprocal Peer Questioning employs no other comprehension strategies.

Adapted from question stems developed by Ryan (1971), King's generic question stems were designed in accordance with the principles of

constructivism and connectionist processing theories. Their primary purpose is to facilitate the encoding and retrieval of new information by activating prior knowledge and personal experiences. By employing the question stems, individuals are encouraged to draw inferences about the new information, to take a new perspective on their existing knowledge, to elaborate the new material by adding details, and to generate relationships between the new material and already existing structures. They will assist in reformulating or restructuring knowledge representations in order to inculcate a deeper understanding of new material (Brown & Campione, 1986; Craik & Lockhart, 1972; Graesser, Swammer, Baggett, & Sell, 1996). Moreover, they ensure that internal and external connections are being made within and between modules or long-term memory networks (Anderson, 1983; McClelland et al., 1986; Rumelhart et al., 1986). Sufficient documentation can be found in the literature indicating that elaboration leads to the superior processing of information (Pressley, McDaniel, Turnure, Wood, & Ahmad, 1987; Pressley, Symons, McDaniel, Snyder, & Turnure, 1988; Woloshyn et al., 1994; Wood, Pressley, & Winne, 1990). Such findings suggests that any approach which incorporates any elaborative activity is likely to produce significant results.

While King conducted several experiments using generic question stems, each with successful results (1989, 1990a, 1991a, 1991b, 1992a, 1992b, 1994a, 1994b), there is one which pertains most closely to this discussion. King and Rosenshine (1993) studied fifth grade students learning science lessons. Students were randomly assigned to three conditions (i.e., highly elaborated question stems, signal words, and unguided questioning) and were instructed to work in pairs. Due to previous successes with college students and guided cooperative questioning, King and Rosenshine sought to determine

whether elementary school children could be trained to use generic question stems and whether they would receive the same benefits as their adult counterparts. As pointed out by the researchers, children typically lack the metacognitive awareness to construct and regulate their own strategies thereby requiring guidance in thinking and knowledge construction. By providing a questioning scaffold, absences in cognitive control may be alleviated. On the other hand, it may be that generic question stems are too sophisticated or cumbersome for younger learners. In previous studies, the simpler signal word strategy (i.e., who, what, where, when, and how) has been shown to facilitate children's reading comprehension (Palincsar & Brown, 1984; Lysynchuk et al., 1990) by stimulating the learner to generate a complete question and by providing an appropriate amount of learner autonomy. This simplicity may make signal words easier to use and remember over time.

Thirty-five grade 5 students were randomly assigned to the three learning conditions and then learning dyads. Five lessons were selected for use in the study. Students were trained to differentiate between memory questions (i.e., simple recall and repetition of lesson content) and think-type questions (i.e., explaining concepts or relationships, applying information to new situations, inference making, etc.). They were then instructed on how to use their respective strategies. Only the unguided question students received no question training of any kind. Students were then given strategy prompt cards in each of the three conditions. Prompt cards for the unguided questioning students simply included instructions to discuss the lessons fully and to ask and answer questions with their partners.

A pre-treatment test was administered to all students following the pretreatment lesson and before the beginning of training. This was followed by

four lessons. After the fifth lesson students discussions were tape recorded. Immediately after, a comprehension posttest was administered. The sixth (transfer) session consisted of a presentation of new material, discussion in pairs (again taped), followed by individual testing. The final (retention) session occurred 6 days after the posttest and consisted only of the administration of a retention test on the material covered in the posttest session. This was followed by the construction of a knowledge map on the same material. All tests were designed to assess literal comprehension of the material as well as inferences beyond the material.

Results indicated that those trained to use the generic question stems performed better on the inferential components of the lesson comprehension posttest than those using signal words or unguided questioning. They also retained more of the learned material (both literal and inferential) six days after the posttest. Finally, students in the generic question stem condition constructed more complete and accurate knowledge maps than those in the other two conditions.

Therefore, it appears that generic question stems induced more complete, accurate, and stable mental representation of the material learned. Such a finding was particularly evident in regards to inferencing. At posttest, generic question stem dyads performed scored significantly higher than the control dyads and considerably better than the signal word dyads. Information was also better retained by the generic question stem dyads suggesting stable knowledge structures. Although students in all three conditions recalled material which was explicitly stated in the lesson equally well at posttest, the generic question stem dyads retained the literal material better over time. According to the researchers, this retention indicates more complex and

lasting knowledge structures. They go on to suggest that despite the same initial training on memory and think-type question generation, generic question stems are superior due to their explicitness and provided structure.

So why might generic questions stems be more preferable than unstructured or semi-structured self-generated questions? For the reasons stated above. That is, generic question stems guide the learner to organize information in meaningful ways, to clarify concepts, and to resolve inconsistencies inherent within one's thinking, the thinking of others, and the studied material(s) (Graesser & McMahan, 1993; King, 1990a). More simply, the cognitive and metacognitive processes of the learner are engaged by way of critical thinking, the activation of pertinent prior knowledge, and through comprehension monitoring (King & Rosenshine, 1993). Unfortunately, such processing is unlikely to occur unless the learner is provided with sufficient structure in which to do so (Graesser & McMahan, 1993; Graesser et al., 1996). As discovered by Graesser and McMahan (1993), the mean likelihood of asking a question about anomalies in algebraic word problems, statistics problems, and analytical brain teasers was .34 under forced question asking and only .04 under self-induced question asking. Therefore, unless prompted, we can never assume that cognitive conflict will automatically lead to a deeper level of comprehension.

Rosenshine, Meister and Chapman (1996) published a review of intervention studies in which learners were taught to generate questions as a means of improving comprehension performance. Twenty-six studies were selected for review. Criteria for inclusion was based upon whether pre- or post-reading instruction on question generation was given, if equivalent experimental and control groups were used, and whether transfer post-tests

were given comparing both groups of students on their ability to comprehend novel materials. One aspect of the review was to analyze the studies according to the different procedural prompts used to help students learn self-questioning. The purpose was to ascertain whether different procedural prompts would yield different results. They included signal words, generic question or question stems, main idea, question types and story grammar. The results indicated that signal words, story grammar, and generic question stems obtained the highest overall median effect sizes on experimenter-developed tests: .85 (80th percentile), 1.08 (86th percentile), and 1.12 (87th percentile), respectively. The lowest effect size was associated with main idea questioning. It generated an effect size of .25 (60th percentile). These statistics were computed by calculating the difference between the comprehension tests means of the experimental and control groups and dividing this result by the standard deviation of the control group.

Based upon these results, it appears that the guidance supplied by generic question stems is superior, albeit marginally, to the guidance provided by story grammar and signal words. Story grammar categories require students to begin with a story grammar and use it as a prompt for generating questions about the narratives they have read. Story grammar includes questions about setting, main characters, and goals. Conversely, signal words are characterized as procedural prompts which require students to begin self-generated questions with *who*, *what*, *where*, *when*, *why*, and *how*. Studies employing this strategy have been conducted by Davey and McBride (1986), Lysynchuk et al. (1990), and Palincsar and Brown (1984), each of which has been previously discussed. According to Rosenshine and colleagues, what makes these three approaches successful is that they provide the learner with

direction, they are concrete, and they are easier to teach and to apply than their counterparts.

While the researchers clearly state the merits of all three strategies, they argue that generic question stems are superior as they "appear to allow students to ask deeper, more comprehensive questions than they could have developed using signal words". Moreover, they "promote deeper processing, initiate recall of background knowledge, require integration of prior knowledge, and provide more direction for processing than might be obtained through the use of the more simplified signal words" (p. 200). While these assertions are intuitively feasible, more research is required to verify the saliency of question stems on cognitive processing and comprehension performance.

Calibration of Comprehension

One of the most important aspects of learning from text is the reader's self-assessment of comprehension. The ability to judge whether one has comprehended material has consequences pertaining to time involvement with text, test preparedness, and the development and sustenance of knowledge structures (Glenberg & Epstein, 1985, 1987; Glenberg, Sanocki et al., 1987; Maki & Berry, 1984; Maki et al., 1990; Weaver, 1990). Often referred to as calibration of comprehension, the correlation between a reader's rated confidence in comprehension and his or her subsequent performance on comprehension questions concerning read passages has been researched by a select few (i.e. Glenberg and colleagues and Maki and colleagues). While it has also been referred to as "metamemory for text" (Maki & Berry, 1984; Maki & Swett, 1987), both members of this comprehension monitoring field assess "the correlation between subjective assessments of knowledge gained from reading

and performance on an objective test" (Glenberg et al., 1987; p. 119). The methodology frequently involves asking subjects to read brief passages of text, to rate levels of confidence in one's ability to answer questions concerning the passages, and to then compare such predictions to subsequent test results. The conclusions drawn from the research indicate that the correlation is typically close to zero (Glenberg et al., 1982; Glenberg & Epstein, 1985, 1987; Glenberg, Sanocki et al., 1987; Morris, 1990). However, there is evidence to suggest that when additional processing is induced, the accuracy of calibration can be improved upon (e.g., Glenberg, Sanocki et al., 1987; Maki et al., 1990; Maki & Swett, 1987). Additional processing techniques of particular interest involve making predictions regarding the memorability of contradictory information (Maki et al., 1987), reading and filling-in deleted letters of words versus reading intact text (Maki et al., 1990), and inducing self-generated feedback on one's level of text comprehension (Glenberg, Sanocki et al., 1987).

In regards to contradictory information, Maki and Swett (1987) found that subjects typically predict that they are more likely to recall an inconsistent idea than one which is consistent with text. For this study, two narrative stories consisting of two versions were written; one which included a similar but consistent fact with the remainder of the story and one which included a fact which was somewhat inconsistent. Each story contained four paragraphs. In a 2 x 2 x 2 mixed design, sixty-four psychology students were randomly assigned to immediate and delayed (10 minutes versus 1 week) and consistent versus inconsistent testing conditions as between-subjects variables and type of rating (memory or importance) as a within-subjects variable. All conditions were required to read both stories. The results indicated that recall was better in the

immediate than in the delayed condition as well as in the inconsistent than in the consistent condition. However, there was an interaction effect between story and consistency. Only in the second story were inconsistent idea units better recalled than consistent ideas, this was not the case in the first story. Moreover, the results were somewhat consistent with subject's memory predictions as they typically predicted that they would better remember an inconsistent idea. However, as demonstrated above, such predictions were only accurate for the second story. Therefore, while generalizations from these results are premature, we can hypothesize that contradictory information does induce a restructuring of established and developing schemas. Consequently, newly acquired information is both elaborated upon and accommodated rendering it memorable for recollection purposes.

Maki et al. (1990) investigated whether paragraphs with deleted letters versus intact text produced higher correlations between one's predictions and subsequent test performance. Employing a within-subjects' design, Experiment 1 manipulated intact versus deleted letters as well as questions concerning comprehension ease or test predictions ratings. Four conditions were created. Half of the 80 subjects received text with deleted letters in odd-numbered paragraphs while the remaining subjects had deletions in even-numbered paragraphs. Below each paragraph were rating scales concerning either comprehension ease or accuracy predictions concerning the test. Moreover, half of the 40 subjects were asked to fill-in the deleted letters while the remaining subjects were to simply figure out what the letters would be. Results indicated that paragraphs with deleted letters led to significantly greater recall than recall from intact paragraphs. Further, the difference between recall with deleted letters and intact paragraphs was greater in the fill-in than in the

mentally fill-in condition. As expected, subjects rated the paragraphs with deleted letters as harder to comprehend than those which were intact, but they did not rate them as more likely to produce poorer performance. Therefore, it appears that ease of comprehension was not the basis for prediction. Gammas were then calculated for individual subjects to verify the accuracy of ratings and test performance for each paragraph. Gammas are nonparametric correlations which require ordinal data. Significant gammas were found in all conditions, but the highest related memory predictions to performance for text with deleted letters as opposed to intact text. Furthermore, higher gammas were also associated with better performance on text with deleted letters. The researchers speculated whether such findings involved increased attention to deleted letter paragraphs at the expense of intact paragraphs. Hence, Experiment 2 sought to test this preferential treatment hypothesis.

To diminish preferential treatment, the researchers employed a between-subjects design. Forty subjects were assigned to either deleted or intact groups. The texts were the same as in Experiment 1 except that every paragraph either contained deleted letters or was intact. In the deleted condition, subjects were instructed to fill-in the missing letters while the intact group were instructed to carefully read for the purposes of understanding. Test predictions were made after every paragraph. Results indicated that there were no cued recall differences for either group. Moreover, there were no differences between the groups concerning predicted levels of performance and confidence at the time of the test. Thus, mean ratings were the same at pretest and at posttest. However, gammas were again calculated for pretest prediction ratings and test scores and for posttest confidence judgments and test scores. Higher gammas in the deleted than in the intact condition replicated the findings

of Experiment 1. Even though there were no differences in recall between the conditions and no differences in the mean ratings, the subjects in the deleted condition predicted their performance more accurately than their counterparts.

Therefore, it appears that increased processing enhances calibration as it indicates to the reader what is and what is not known. Perhaps even more interesting, increased processing need not be induced for all paragraphs as the results were similar for both experiments. What undergirds these findings can only be speculated. As suggested by Maki and colleagues, the need to think about incomplete words may coerce the reader to assess his or her knowledge of the material in each paragraph more carefully than when the material is simply read. Moreover, it may solicit the vital self-generated feedback necessary for metacomprehension.

The vitality of self-generated feedback was demonstrated by Glenberg et al. (1987). They hypothesized that a significant difference between an accurate and inaccurate metacognitive judgment is a function of self-generated feedback. During calibration of comprehension experiments, subjects read a text, predict performance on a test, and then take the test. The feedback derived from the test results follows a subject's comprehension prediction thereby making it too late to accurately determine whether information was or was not acquired. To combat this design flaw, Glenberg and colleagues conducted three experiments which included pretests. On the basis of feedback from these tests, subjects accurately predicted future performance on the same items. However, as suggested, the same items appeared on both the pretest and posttest subsequently negating the naturalistic practicality of such feedback. After all, it is unfeasible to retest students on the same material. Nevertheless, some applications can be considered. Most notable for this discussion,

Glenberg and colleagues suggest that more connections should be made between the text and the test. This includes advanced organizers and textual signals. Also worthy of consideration, although not suggested, are generic question stems. It is possible that they would induce the reader to consciously generate connections between ideas within the text as well as to prior knowledge structures. Such generation would provide notification of comprehension weaknesses and poor knowledge acquisition. Thus, calibration of comprehension may be enhanced with self-generated feedback resulting in improved knowledge acquisition, utilization, and maintenance.

While it is of obvious importance to understand how calibration of comprehension can be improved upon, we must also understand why calibration tends to be so poor. Research has been conducted to determine those variables which undermine correlations between prediction and test performance. For instance, in a majority of studies the text being studied is expository. The utilization of expository text closely approximates a typical reading experience for university students suggesting that familiarity with a text genre may incite over confidence. The employment of expository text also provides the opportunity to determine whether knowledge of, and/or expertise with text content influences calibration of comprehension.

Glenberg and Epstein (1987) investigated the salience of high versus low domain knowledge on calibration. Based upon their own misgivings regarding previous research results, Glenberg and Epstein queried whether a lack of domain knowledge induces poor calibration outcomes. They argued that we intuitively surmise that domain expertise enables a reader to accurately predict his or her performance following the study of domain specific material. Conversely, performance regarding material outside one's field of expertise

may be more difficult to predict. Despite the merit of such reasoning, its fallibility is pointed out (e.g., Bradley, 1981; Hock, 1985; Oskamp, 1965)(as cited in Glenberg & Epstein, 1987, p. 85).

Glenberg and Epstein (1987) recruited 70 subjects who had completed two university courses in either physics or music theory. Thirty-two pieces of text were divided between 2 booklets; one for each testing session. Eight were music texts alternated with 8 physics texts. Following the text passages in each booklet were 16 sets of 5 probes. Each set corresponded to each piece of text, and the sets were in the same order as the texts. Probe 1 was a confidence probe requiring the subject to indicate confidence in his or her ability to judge the correctness of an inference regarding a reference to the central principle. Probe 2 was an inference test asking the subject to judge the correctness of a textual inference located on the following page. Probe 3 was a confidence scale. Each subject was asked to rate his or her confidence of a correct response to the inference question. Probe 4 was a recalibration of confidence scale requiring the subject to indicate confidence in his or her ability to answer another inference. Finally, Probe 5 was the second inference test. For this test, Glenberg and Epstein hypothesized that the first inference question would provide valid cues to the degree of one's comprehension and would in turn improve his or her ability to predict future performance.

Tested in small groups, the subjects were allowed to read the text at their own pace. The first session consisted of the 16 pieces of text and subsequent probes. The second session was scheduled for 1 to 7 days later. At the end of the second session, the subjects completed two questionnaires. Allowed to refer back to the booklets, they were asked to rate topic familiarity concerning all 32 pieces of text on a scale ranging from 1(very low) to 6 (very high). The

subjects were then surveyed regarding their domain-specific experiences and dualistic tendencies. That is, whether they adhere to dualistic principles (i.e. that truth is absolute in most if not all domains) or relativistic principles (i.e. that truth is determined by the context and that propositions are true or false within a particular frame of reference). According to Ryan (1984), relativists engage in more sophisticated comprehension monitoring than dualists.

Using the Goodman-Kruskal gamma (G) to measure calibration, the researchers determined that calibration and recalibration declined with domain expertise. Such a decline was significant for students' knowledgeable in physics. Referring to the self-classification hypothesis, it appeared that the subjects were not assessing knowledge gained from a particular text; they were responding according to their beliefs regarding ability within a given domain. Thus, confidence ratings were decided according to self-classification and not assessment of text comprehension. This same self-classification strategy appeared to be reapplied when the subjects were asked to reassess their confidence in future performance. The results suggest that subjects do not take advantage of experience gained while answering an inference test to predict performance on future tests. Finally, across domains, the results were markedly different as the subjects were calibrated. Subjects unfamiliar with a specific domain not only judged themselves as less confident, they were more likely to be incorrect when answering inference questions.

Similar results were replicated by Glenberg, Sanocki et al. (1987). Using confidence and inference verification procedures, Experiment 3 demonstrated that domain familiarity judgments are highly correlated with confidence judgments supporting the claim that confidence is based upon familiarity. Because the results did not distinguish between effects of domain familiarity

and effects of familiarity with particular texts, Experiments 4 and 5 were conducted. By manipulating the central principle of the text with either a paraphrase or a verbatim statement, the researchers were able to demonstrate in Experiment 4 that this manipulation does affect familiarity with particular statements from the text. Repeated again in Experiment 5, it was further shown that the verbatim and paraphrase conditions did not differ significantly with respect to confidence or performance on the inference test. Thus, familiarity with particular statements does not control confidence. As pointed out by Glenberg and colleagues, the domain familiarity hypothesis appears to satisfy these results, but only by default.

The purpose in discussing these studies is to demonstrate that unfamiliarity with a domain of study can facilitate calibration of text comprehension. Why that is the case is unclear. However, it does appear that readers typically base their judgments solely upon domain familiarity. While it is unsafe to assume that high or low domain familiarity necessarily leads to poor or superior calibration, steps must be taken to alleviate potential ill-effects. Once again, the utilization of generic question stems may serve a strategic purpose by alleviating the over-confidence associated with high domain familiarity. That is, question stems may inform the reader of his or her level of comprehension by way of higher level thinking and querying. Such knowledge will enable the reader to make informed and self-controlled comprehension repairs.

Another factor concerning poor calibration may involve text structure. That is, if the text is coherent versus incoherent, will calibration of comprehension be affected? According to Weaver, Bryant, and Burns (1995), virtually no one has looked at how text revisions can influence metamemory for

text. In an unpublished master's thesis by Burns (1993) (as cited in Weaver et al., 1995), an Air Force textbook was revised in several ways: through a *principled version* (repairing text at those locations in which the reader would have to make an inference to establish coherence); a *heuristic version* (performed by an expert in text revision to make the text as understandable as possible); and a *readability version* (designed to increase the object readability of the passage to match that of the heuristic version; achieved by shortening sentences and including more frequently used words). In Experiment 1, only the original, heuristic and readability versions were used. Results indicated that the heuristic version was superior on three dependent variables: reading time, confidence ratings, and performance on a multiple-choice test. Readers spent 10% less time reading the heuristic version, they rated their level of comprehension higher, and they performed better on the multiple choice test. Using the gamma correlation, the last two variables were correlated. Readers of the original and readability versions produced correlations that were not statistically different from zero. Conversely, readers of the heuristic version displayed gammas of .37. Thus, the more understandable the text, the better one's predicted and actual comprehension are calibrated.

Experiment 2 included the principled version. Because heuristic revisions generated such gains in comprehension monitoring, Burns wanted to see whether they could be replicated with a different type of revision. Procedures and variables remained the same as in Experiment 1. The results were rather puzzling for the researcher and advisory committee as the heuristic version proved to be superior in confidence ratings and comprehension monitoring. The correlation between predicted and actual performance for the principled revision was only .18, which was not statistically different from zero.

Furthermore, the correlation for the readability version was as high as for the heuristic version (.30). Hypotheses for these results included the domain familiarity hypothesis (Glenberg et al., 1987) (i.e., that familiarity with a domain can generate a false sense of comprehension) and a failure to engage in active or deep processing (i.e., the text was too readable causing readers to skip important details). Neither, however, appealed to the researcher and his advisory committee. They remained stumped by the results.

Weaver et al. (1995) believe that some answers may be found to this puzzle in a study conducted by Weaver and Bryant (1994) (as cited in Weaver et al., 1995). Instead of using exclusively narrative or expository text selections, as is the case in all other calibration of comprehension research, Weaver and Bryant decided to compare both genres in regards to confidence ratings and recognition test performance. They felt that some of the inconsistencies within the literature are due to differences in text selections and the information which can be subsequently garnered. Typically, narrative text warrants thematic based processing while expository text induces detail-oriented processing. Thus, a multiple choice performance test was designed and written to ascertain relational information (i.e., theme) and interitem information (i.e., detail).

The results indicated that subjects consistently performed better on the questions from the narrative text and rated it as the easiest to read. However, the gamma correlations between predicted and actual performance demonstrated that neither group did consistently poorer when comparing recognition performance of relational and interitem tasks. However, readers of the narrative text were better able to monitor their performance on relation questions and readers of expository text better monitored interitem performance. Therefore, depending upon the types of questions asked by the

researchers (i.e., relational versus interitem) and/or the text genre, calibration of comprehension can be affected.

Weaver and Bryant (1994) also noted that age-appropriateness may have an impact on predicted and actual performance. The narrative text selections were comprised of fairy tales for children. Therefore, in Experiment 2 they constructed a whole new set of stimuli. Three reading levels were determined for the narrative and expository selections: easy (below grade 8); standard (around grade 12); and difficult (four years beyond grade 12). Recognition performance followed this pattern. Each subject read 2 narrative and 2 expository passages of the same difficulty level.

Results indicated that those who read the easiest text scored the highest followed by readers of the standard text and readers of the difficult text. This same linear trend was seen in the confidence ratings. However, the gamma correlations indicated that the most accurate predictions with performance were amongst readers of the standard text. This suggests that calibration of comprehension can be improved upon when text is adjusted to the appropriate level of readability. In the earlier Burns (1993) experiments, it may be that the text was neither age-appropriate or well-matched between text genre and test items.

Therefore, domain familiarity, text structure, text genre, and age-appropriateness can all uniquely affect calibration. Controlling for these variables can be problematic given the diversity of today's classrooms. Once again, the utilization of generic question stems may alleviate volatile combinations by signaling comprehension breakdowns. Research in this area may prove to be informative.

Note-taking, Written Recording, and Memory

Note-taking has long been a means of encoding and storing verbal and written sources of information. The processes involved in note-taking include the processing and reprocessing of textual and verbal materials. As demonstrated by Peper and Mayer (1986), note-taking is an encoding process which helps learners to build connections not only between the informational units contained within the source being studied, but to one's prior knowledge. In regards to the process of storing such information, by reviewing self-generated and instructor-provided notes, the adaptive cognitive processes of organization and elaboration may be elicited (Kiewra, 1988, 1989; Spires, 1993).

Because of the potentially meritorious outcomes of appropriate note-taking, its encouragement prior to explicit strategic instruction is ill-advised. According to Kiewra (1988, 1989), Peper and Mayer (1986), Kraker (1993), and Spires (1993), note-taking is a specific strategy which must be learned in order to be effective.

Kiewra (1989) suggests that notes can be made effective in three ways "...they can be made more complete (e.g., Fisher and Harris, 1973); they can specify internal connections or relationships among existing lecture ideas (e.g., Kiewra et al., 1988a); and they can connect lecture information to previously acquired knowledge (e.g., Peper and Mayer, 1986)" (p. 158). In all three cases, he substantiated the validity of his suggestions with research findings. In an earlier publication, Kiewra (1988) detailed the constituents of an effective strategy training program. They include an "... (a) awareness of the significance of the strategy, (b) explicit knowledge of when and how to use the strategy, and (c) training in monitoring the application of the strategy" (p. 47). Therefore,

subjects must be imbued with cognitive and metacognitive awareness in order to be appropriately trained.

The need to instruct students on note-taking strategies has also been argued by Kraker (1993). In a study designed to examine the written notation of normally achieving and learning disabled grade one students, Kraker confirmed earlier reports that learning disabled students lack organizational strategies, that they have difficulty comprehending and processing auditory information, and that they exhibit problems with spelling and handwriting; each being a component of note-taking. It is her suggestion that explicit guidance be given through such written language competencies as note-taking (graphic and linguistic), graphs, charts, and text production. These tasks will ameliorate the restrictive effects of memory deficits, attentional problems, and mechanical limitations.

In support of explicit instruction, Spires (1993) argued that metacognitive awareness will arise when the purposes for training are articulated and when strategic modeling, scaffolding, and feedback are provided. He hypothesized that self-monitoring processes such as reviewing (i.e., external storage function) are salient information processors (e.g., Kiewra, 1989) as they correlate with high achievement levels. Any activity which by-passes self-monitoring, such as the singular act of note-taking (i.e., an encoding function), will therefore be inadequate. Spires (1993) also cited King's (1989) investigations of self-questioning training as being cognitively and metacognitively enhancing. King's findings suggest that posing questions of a cognitive and metacognitive nature leads to improved lecture comprehension. Therefore, Spires sought to examine the effects of explicit note-taking instruction, with and without a comprehension monitoring activity (i.e., self-questioning), on the quality of note-

taking and on the immediate and delayed comprehension of lecture information.

Ninety-nine college freshmen were divided into three treatment groups: explicit note-taking instruction, explicit note-taking instruction with self-questioning, and a control group that received no explicit instruction. Similar to the explicit note-taking condition, the students in the self-questioning condition were instructed in the use of the split-page method of note-taking. This approach requires the note-taker to write the main ideas on the left side of the page with corresponding supporting details (i.e., definitions, examples) on the right side. Students were also informed on the usefulness of the strategy, they were instructed to observe instructor modeling, and they were given feedback on the quality of their notes. Eventually they were left to work independently. Unlike their counterparts, the self-questioning subjects were also taught to monitor their note-taking. They were instructed on how to query their planning (e.g., What is my purpose in listening to this lecture?), monitoring (e.g., Am I maintaining a satisfactory level of concentration?), and evaluation (e.g., Did I deal with comprehension failures adequately?) knowledge, skills, and strategies.

Results indicated that the self-questioning strategy leads to superior note quality and the immediate comprehension of lecture information. Spires reasons that the strategy assists students to mentally prepare to take notes, to monitor their levels of comprehension during note-taking, and to evaluate how well they took notes after the lecture was over. More simply, this strategy requires students to address potential difficulties during comprehension.

These findings appear promising for programs which combine note-taking and self-questioning during lecture format. Whether such findings

transfer to reading comprehension requires further research. Moreover, in keeping with the interests of this discussion, it is important to determine whether similar findings will manifest under the combination of self-generated questions, which are created from generic question stems and which serve as skeletal outlines for notes and for reading comprehension. After all, generic question stems not only induce external and internal connections (King, 1990a), they guide or cue the subject to make note of important details. Therefore, generic question stems may elicit effective note-taking skills, lead to a complete set of notes available for further review, induce elaborative/generative processing skills, and help develop metacognitive awareness.

As indicated, note-taking should not be depicted as the simple act of recording. To qualify as an adaptive means of processing and encoding information, explicit instruction on appropriate procedures and strategies is required. However, there is some suggestion that the act of recording has mnemonic benefits (Luria, 1978). Vygotsky (1986, 1987) was a forerunner in suggesting the connection between writing and memory. He suggested that written speech is the key to inner speech since it translates that which is conscious and intellectual. Luria (1978) expanded upon this philosophy when he instructed 3 to 5 year old children to use writing as a means of recall. The children were to record and to later review their notations of dictated sentences. According to Luria, because the children were able to accurately recall, written recording inherently contains mnemonic benefits. Olson (1994, 1996) concurs with this reasoning by reiterating the links between consciousness and the written word. He writes that the evolution of written language has preserved the cultural and historical integrity of most societies by extending the memory of their members. While Olson does not directly refer to the simple act of

recording, we can assume that transposing a word into print evokes, at least in part, the conscious processing to which he refers.

According to the levels of processing theory (Craik, & Lockhart, 1972), how well an item will be remembered is contingent upon how deeply it was processed. The deeper the analysis, the more persistent the memory trace. In their seminal work on levels of processing, Craik and Lockhart (1972) write, "retention is a function of depth, and various factors, such as the amount of attention devoted to a stimulus, its compatibility with the analyzing structures, and the processing time available, will determine the depth to which it is processed" (p. 676). They go on to add that stimuli can be retained by "keeping the items in consciousness" (p. 676). In conjunction with the theories of the above-cited scholars, it can be argued that requiring one to record his or her questions and responses invokes deep processing since it requires attention, conscious processing, and an understanding of language.

CHAPTER THREE

METHOD

The purpose of this experiment was to examine whether "questioning with generic question stems", "questioning using signal words", or "unguided questioning" influence comprehension of expository text which has either good local coherence (C) and good macrostructure (M) (i.e., high coherence; CM) or poor local coherence (c) and poor macrostructure (m) (i.e., low coherence; cm) when participants possess either high domain knowledge or low domain knowledge. Outcomes were measured using an immediate post-test comprised of textbase and situation model questions (textbase, bridging inference, elaborative inference, and problem solving); a summary recall question to be scored for text propositions recalled, inference items, and macrostructure levels analysis; and a pre- and post-reading sorting task designed to measure changes in the reader's conceptual structure of text content. A confidence measure was administered to determine calibration of comprehension and levels of reader confidence.

Participants

A domain knowledge questionnaire (see Appendix B) was given to 193 participants in Psychology 110. Of those participants, 80 indicated that they would be interested in partaking in the second phase of the study (i.e., the experimental portion). The questionnaires of those 80 willing participants were

then scored. A median-split procedure on the scores was employed. The median score was 20.9 with a range of 6 to 55 out of a possible total of 73. This procedure distinguished scores as being indicative of high or low domain knowledge. The mean of the high domain knowledge participants was 37 (S.D. = 9.0) while the mean of the low domain knowledge participants was 12.4 (S.D. = 4.3). On the basis of their individual scores, participants were then sorted on the basis of high or low domain knowledge. They were then randomly assigned to treatment conditions. That is, low domain knowledge participants reading high coherence text (CM) with either generic question stems, signal words, or unguided questioning strategies; low domain knowledge participants reading low coherence text (cm) with either generic question stems, signal words, or unguided questioning strategies; high domain knowledge participants reading high coherence text (CM) with either generic question stems, signal words, or unguided questioning strategies; or high domain knowledge participants reading low coherence text (cm) with either generic question stems, signal words, or unguided questioning strategies.

The researcher then contacted the 80 participants asking whether they would be willing to continue with the second phase of the experiment and appointments for group testing were set. Sixty-eight individuals were scheduled to partake in the experiment. Only 63 individuals participated due to absences. This led to unequal cell sizes for analyses.

Research design

The design of the experiment is a 3 x 2 x 2 (Treatment x Text Coherence x Domain Knowledge) factorial MANOVA.

Materials

Training materials were compiled for the second phase of the study. They included training scripts, practice texts, and questioning instructions (see Appendix C). Following training, strategy prompt cards (Appendix D) were given to the participants in the generic question stem condition, the signal word condition, and the unguided questioning condition. The experimental texts were acquired from Eileen Kintsch by way of personal correspondence (Appendix E and Appendix F). They were based upon an entry in a science encyclopedia for school-age students (Raintree Illustrated Science Encyclopedia, 1984). Two versions of the same text were used (e.g., maximum coherence at the local level and maximum coherence at the macrolevel; minimum coherence at the local level and *minimal* coherence at the macrolevel). Local coherence was maximized by McNamara et al. (1996) through the following revisions:

- “1. Replacing pronouns with noun phrases when the referent was potentially ambiguous (e.g., replacing *it* with *the heart*).
2. Adding descriptive elaborations that link unfamiliar concepts with familiar ones (e.g., ‘This disease usually follows a sore throat caused by bacteria known as streptococci. *This is often called strep throat.*’).
3. Adding sentence connectives (e.g., *however, therefore, because, so that*) to specify the relation between sentences or ideas.
4. Replacing words to increase argument overlap (e.g., replacing *person* and *cases* with *baby* or *babies*).” (McNamara et al., 1996, p. 21).

Revisions to global coherence, or the macrostructure of the text involved emphasizing macropropositions. To attentive and knowledgeable readers, these same macropropositions could be inferred in the minimally coherent

macrostructure text. The means of explicitly signaling the macropropositions included: "a) adding topic headers (e.g., *congenital heart disease, acquired heart disease*) and b) adding macropropositions serving to link each paragraph to the rest of the text and overall topic (e.g., 'There are many kinds of heart disease, some of which are present at birth and some of which are acquired later.')

 (McNamara et al., 1996, p. 21).

The supplemental text entitled "Blood" was written by the researcher (Appendix G). An attempt to match syntax, semantics, explanatory patterns, and readability levels were carefully pursued (see Table 1). Moreover, to ensure that the readability formula of the experimental and supplementary texts were appropriately matched, comparisons were made with five, randomly selected paragraphs, each taken from two introductory psychology textbooks. It was concluded that the experimental texts were of an appropriate reading level for the participants. The text was neither too easy nor too difficult. Hence, comprehension was deemed feasible.

Each group received the text entitled "Blood" prior to the experimental text. This text was written to complement the "Heart Disease" text so that links may be drawn during the strategic questioning portion of the experiment. Such supplementary information also facilitates the development of a situation model by enabling a broader understanding of the experimental topic and by providing some prior knowledge to which new information may be linked.

A domain knowledge questionnaire (see Appendix B) was given during the screening procedure. For the second phase of the experiment, a pre-sorting task was given prior to any reading (see Appendix H). Following the experimental phase of this study, students were given a calibration questionnaire (see Appendix I) which was matched to 3 questions from each

Table 1

Readability Scores

Readability Index*	Blood	Heart Disease (cm)	Heart Disease (CM)	Psych 1**	Psych 2***	
Flesch Reading Ease		60.28	69.10	68.90	52.80	56.18
Flesch-Kincaid GL		8.93	7.40	7.10	9.77	9.70
Coleman-Liau GL		14.26	9.87	9.93	11.30	14.15
Bormuth GL		10.90	9.50	9.50	10.30	10.60

* Microsoft Word 6.0.1

**Baron, R.A., Earhard, B., and Ozier, M. (1998) Psychology (2nd Canadian ed.). Toronto: Allyn and Bacon.

***Carlson, N.R. and Buskit, W. (1997) Psychology: The science of behavior (5th ed.). Toronto: Allyn and Bacon.

section of the post-test. Using the same post test design as McNamara, et al. (1996), the post-test is comprised of textbase, elaborative inferencing, bridging inferencing, and problem-solving questions. "Textbase questions can be answered on the basis of the text base; elaborative inference questions require, in addition, some outside knowledge but not a very specific situation model; answers to bridging-inference and problem solving questions, in contrast, depend on a well-formed situation model" (McNamara et al., 1996, p. 22). Therefore, the post-test provided data on both text memory and text learning (see Appendix J). There were 3 questions representing each question type. Like McNamara et al. (1996), questions within each set were matched as well as possible for difficulty and content matter.

The post-sorting task is identical to the pre-sorting so as to gauge changes in situational conceptions. Finally, a summary recall (see Appendix K) was required to ascertain details recalled from the text as well as perceptions of macrostructure. This too helped to clarify whether treatment, domain knowledge and/or text coherency significantly impact text memory and text learning.

Instruments and analysis

Prior knowledge questionnaire. Participants were given a prior knowledge questionnaire to ascertain levels of knowledge and understanding of the heart. The first section of the instrument called for information regarding previous course work (i.e., high school classes, university classes, and emergency medicine courses) and personal relationships with those diagnosed with heart disease. The second section included a diagram of the human heart. Participants were asked to label as many parts as possible. Guessing was

encouraged so that implicit knowledge may be activated and retrieved. The final section included 10 multiple choice questions which assessed basic knowledge of the heart, its functions, and potential disorders. The questionnaire was scored out of a maximum of 73 points. For a more detailed description of the scoring criterion see Appendix B. Scores for all of the participants who agreed to participate in the second phase were tallied. On the basis of a median-split procedure, participants were randomly assigned to high or low domain knowledge categories.

Post-test. Participants were given a post-test comprised of 12 short answer questions (see Appendix J). The purpose was to assess the quantity, quality, and type of knowledge acquired from the heart disease text. Therefore, both textbase and situation model representations were assessed by way of the previously discussed question types (textbase, bridging inference, elaborative inference, and problem solving). The textbase model was represented by textbase questions as they are answered on the basis of the textbase. The situation model was represented by bridging inference and problem solving questions. Elaborative inference questions are indicative of both textbase and situation models as they require prior knowledge, but not a very specific situation model.

Each question type appeared three times. The number of points awarded to each question ranged from 0 to 7. Item difficulty was set reasonably high to avoid the possibility of a ceiling effect. For a detailed description of the scoring criterion and the post-test questions and answers, see Appendix J. An ANOVA was conducted on the post-test total score while a MANOVA was conducted on the question types using the factors Treatment x Text Coherence

x Domain Knowledge.

Summary recall. Participants were instructed to write a summary of the heart disease text (see Appendix K). They were asked to “state the important information in the text in a briefer form”. Because a summary recall is comprised of reproductive and reconstructive components, a mixture of reconstructions derived from a reader’s situation model was anticipated. This is especially the case when an individual is high in domain knowledge or when the textbase itself can no longer be successfully retrieved. For those participants with low domain knowledge, or in the case of short-term experiments, a textbase reproduction can be expected (McNamara et al., 1996). Thus, this task quantified a reader’s textbase memory in addition to revealing situational reconstructions.

Summaries were scored for a) the number of text propositions (i.e., both texts share 65 propositions and 7 macropropositions)(see Appendix K) included in the summary/recall and b) the inclusion of nontext items. Nontext items were propositionalized and assigned to the following inference categories: 1)

Generalizations: These are reductive inferences which are constructed from more detailed statements in the text. They can be traced to the actual propositions they subsume, with the exception of global generalizations, which are inferences about the overall meaning of the text. Generalizations typically reduce the number of propositions within the text by a minimum of one (E.

Kintsch, 1990); 2)***Elaborations:*** These are inferences which are not directly implied by the text. Rather, they originate from the subject’s own knowledge about the content of the text or any related information (E. Kintsch, 1990); and 3)

Reorderings: These are inferences which rearrange text content in an order

which differs from the original text. Only between paragraph reorderings were considered. They were not scored at the propositional level. Rather, an idea in the form of a sentence or paragraph was considered a reordering if it required backtracking to an earlier part of the text (E. Kintsch, 1990). Thus, a textbase model representation is revealed by way of propositions and a situation model representation is comprised of generalizations. Elaborations and reorderings are indicative of a less well-developed situation model.

A liberal gist scoring criterion was used. For every proposition, generalization, elaboration, and reordering which appeared, 1 point was awarded. Therefore, an ANOVA was conducted on the text propositions while a MANOVA was conducted on the generalizations, elaborations, and reorderings using the factors Treatment x Text Coherence x Domain Knowledge.

A macrostructure levels analysis was also conducted. It is necessary to determine whether participants are following the macrostructure of the experimental text or whether they are developing their own macrostructure. Such an investigation is worth undertaking since the construction of macrostructure is necessary for situation model development and for a deeper understanding of text meaning.

In preparing to measure the levels of macrostructure, it is necessary to note that macrostructure "consists not only of generalizations of textual details but also of propositions selected from the text on the basis of their importance to the overall meaning. Such propositions also function as macropropositions. Therefore, it is important both to examine the amount of generalized information in a summary and to see if the information included is macrorelevant" (E. Kintsch, 1990, p. 167). Thus, the experimental text was used as a scoring template (see Appendix K) and points were awarded for the number of

statements mentioned at each level. Levels of importance are as follows: *Level 1* consists of topic statements (i.e., a label or a more elaborated topic statement); *Level 2* is composed of inferred subtopics; *Level 3* consists of other text-based macropropositions that function as subheadings for groups of detailed statements. Some are inferred and some are mentioned in the text; and *Level 4* comprises a representative but not comprehensive list of concrete details from the text. Again, 1 point was awarded for each detail mentioned for each level. The scores were then converted to percentages. As with the summary recall inference analysis, a textbase model representation was revealed by way of Level 4 citations while Levels 1, 2, and 3 citations are indicative of a less well-developed situation model representation.

A MANOVA was conducted on the 4 levels cited above using the factors Treatment x Text Coherence x Domain Knowledge.

Sorting task. The inclusion of a sorting task has been recommended by Kintsch (1998; McNamara & Kintsch, 1996; McNamara et al., 1996) as an effective means of determining changes in a reader's conceptual structure and the amount learned as a result of reading a text. More simply, it is an appropriate method for assessing situation model understanding. If the text has an effect on the reader's memory and knowledge, changes will arise in the way the reader organizes a knowledge domain, and the changes will be in the direction of text organization (Kintsch, 1998). Thus, according to Kintsch, the focus of the task is not on how well the subjects sort the items, but in the degree to which the information in the text influences their sorting (McNamara et al., 1996).

Participants were given 18 concept words to sort before and after

reading the experimental text. They were instructed to put the words into categories according to how they thought the concepts should go together. This was completed on the same piece of paper on which the concept words were given. Participants were told that they can make as few or as many categories as they wish; that there are no restrictions or limitations as to how many words they can put into each group; they can change their minds and reorganize the categories at any time; and that there is no correct or incorrect way to organize the words. At the end of each category they were asked to briefly state why they put the words together in the manner that they did and how they would label each category (see Appendix H).

The 18 concept words were selected by McNamara et al. (1996). There are eight nontext items: three refer to parts of the human heart (ventricle, mitral valve, pulmonary vein); two are body organs (thyroid, kidney); three are related to diseases not specific to the heart (cancer, multiple sclerosis, malignant). The remaining 10 items are directly from the text: three concepts relate to congenital heart disease (blue baby, septal defect, carbon dioxide); and seven concepts relate to acquired heart disease (rheumatic fever, streptococci, blood clot, coronary thrombosis, by-pass surgery, arrhythmia, pacemaker). The motivation for selecting those items was to "provide a group of concepts for which there were not only several rational sorting principles, but also clearly discernible, text-driven sorting principles" (McNamara et al., 1996, p. 22). Because these are the same words used by McNamara et al. (1996) during their text comprehension experiment, categories were scored in the same manner.

In establishing the sorting score, McNamara et al. (1996) state that it "is a measure of harmony between the participant's sorting matrix and a weight matrix indicative of an ideal sorting. Specifically, the sorting score is the sum of

the inner product between the participant's sorting matrix and an ideal sorting matrix. This sum is divided by the total of the positive values in the matrix, in this case 25 (i.e., 12.5×2), so that the sort score varied between +1.0 and -1.0¹ (McNamara et al, p. 28-30)¹. For the ideal sorting matrix, the researchers assumed that the most credit should be given when participants sort into the same category those items that were closely related in the text. Specifically, the following five sets of items were assigned a weight of 1: (a) blue baby, septal defect, and carbon dioxide; (b) rheumatic fever and streptococci; (c) blood clot, coronary thrombosis, and bypass surgery; (d) arrhythmia and pacemaker; and (e) bypass surgery and pacemaker. Less credit (a weight of .5) was given for nontext items that were correctly categorized: (f) ventricle, mitral valve, and pulmonary vein; (g) thyroid and kidney; and (h) cancer, multiple sclerosis, and malignant. All other sorts were assigned a negative value (-.0456204) in the ideal sorting matrix so that the sum of the weight matrix was zero (see Appendix H).

A repeated measures analysis was performed with the between-subjects variables being Treatment x Text Coherence x Domain Knowledge and the

¹By multiplying two matrices or vectors, that is, calculating the inner products of the two, one derives a statistical computation of the similarity, or harmony, of the two matrices. Thus, the sort score is the sum of the inner product between a participant's sorting matrix (s_{ij}) and a weight matrix (w_{ij}) divided by the sum of the positive weights in the weight matrix ($\sum w_{ij}/2$): $((\sum (s_{ij} \times w_{ij})) / (\sum w_{ij}/2))$. Specifically, a participant's sorting matrix consists of 1s and 0s, whereby a 1 indicates a pair of items sorted together and a 0 indicates a pair not sorted together. The weight matrix, indicative of an ideal sorting, was constructed by assigning a weight ($w_{ij} = 1$) to cells of the matrix representing pairings between items closely related in the text. A weight ($w_{ij} = .05$) was assigned to pairings between nontext items which were correctly categorized. This yields a total of 25 points for the positive values in the matrix (i.e., 12.5×2). The remaining cells of the weight matrix were assigned negative values in such a way that the sum of the matrix became 0. Specifically, the negative weight was set equal to $-(\sum w_{ij}/2)/(m - k)$, where $\sum w_{ij}/2$ is the sum of the positive values in the matrix, m is the total number of nondiagonal cells in the matrix $n^2 - n$, and k is the number of cells containing positive values. Diagonal cells of the weight matrix were assigned a weight of 0.

within-subjects variable being pre-treatment sorting scores and post-treatment sorting scores.

Calibration of comprehension. For this study, calibration of comprehension is the correlation between a participant's confidence rating on how well he or she knows information derived from the text and his or her test performance on specific questions regarding that information. Separate calibration coefficients were computed for each participant by measuring the association between his or her confidence ratings and performance scores on matched, open-ended questions. Confidence ratings were measured on a Likert scale ranging from 1(not at all confident) to 7 (very confident) (see Appendix I) while performance scores on matched questions were scored from 0 to 7.

The two sets of scores were used in computing the Goodman-Kruskal gamma correlation (G) (Goodman & Kruskal, 1954, 1959, 1963, 1979) (as cited in Nelson, 1984; Glenberg & Epstein, 1987; Maki, Foley et al., 1990). Gamma is a nonparametric correlation which requires ordinal data. According to Nelson (1984), it is the most appropriate index of association for measuring this type of metacognitive performance data. Like the Pearson product-moment coefficient, G ranges from -1 to 1, with zero indicating no relation. Therefore, Pearson product-moment coefficients were also calculated and reported.

Gamma correlations and pearson product moment coefficients were computed for each participant, correlating the confidence with the proportion correct across all 12 post-test and confidence questions. A MANOVA was then conducted on the correlation coefficients using the factors Treatment x Text Coherence x Domain Knowledge. An ANOVA was also conducted on the

confidence scores using the factors indicated above to determine which factor(s) affect one's level of confidence.

Procedure

Permission was sought and granted by the University Advisory Committee on Ethics in Behavioural Science Research (see Appendix L). This was followed by the screening procedure. It was led by the researcher and two other research assistants. Participants were given the domain knowledge questionnaire (see Appendix B) as well as a consent form, a detailed description of the purpose of the questionnaire, and an overview of their rights and privileges as a participant (see Appendix M).

Following the screening procedure, questionnaires were sorted on the basis of those consenting to participate in the second phase of the study. The questionnaires were scored and participants were randomly assigned to treatment conditions on the basis of their scores. The researcher then contacted the participants asking whether they would be willing to continue with the second phase of the experiment and appointments for group testing were set. Participants could select the most appropriate time slot from an option of two periods. Thus, group sizes for the experiments ranged from 6 to 13 participants. If either of the 2 periods were inappropriate, some accommodations were made in the form of additional testing times or switching to another treatment condition which held a more suitable time slot.

During the second phase of the experiment, participants were reminded that the purpose of the study is to understand how people read and comprehend text. They were then asked to sign the consent forms and were given an information sheet regarding the intentions of the experiment (see

Appendix N).

Following the general introduction to the experiment, participants in the signal word and generic question stem conditions were instructed on the differences between memory questions and critical thinking questions. That is, they were trained to differentiate between memory (textbase model) questions and critical thinking (situation model) questions. Memory questions simply require participants to recall the information within the text or information that is memorized. Critical thinking questions require participants to explain concepts or relationships, link ideas within the text, apply information to new situations, make inferences, provide justifications, and so forth (King, 1990a; King & Rosenshine, 1993).

Participants in the signal word and generic question stem conditions were then instructed on the fundamentals of their questioning condition. They received strategy prompt cards and were shown how to read expository text and develop the types of questions discussed above using either signal words or generic question stems. Participants in the unguided questioning condition did not receive any training in question generation (King, 1991a; King & Rosenshine, 1993). They were simply told to develop questions which are based on the text(s). The participants received a prompt card telling them to create the best questions that they can think of regarding the text. Participants were asked to generate 6 questions. In total, this training period took 20 minutes. However, for the unguided questioning group, the additional training on question development and practice was unnecessary. Therefore, they immediately moved on to the pre-reading sorting task.

Following the practice session, the generic question stem and signal word conditions were then given the pre-reading sorting task. Participants were

told to put the words into any category they saw fit and that there are not correct or incorrect ways to make categories. Furthermore, they could use a word more than once. Participants were given exactly 10 minutes.

Following the pre-reading sorting task, all participants received the text entitled "Blood" to read silently and individually. Participants were given exactly 5 minutes. The purpose in reading this text was to provide supplementary information so that external connections or elaborations could be made with the experimental text entitled "Heart Disease". That is, the "Blood" text is not specifically related to the experimental text. Rather, it provides additional information which may link to existing domain knowledge or enhance the depth of new information. As indicated above, participants in the generic question stem and signal word conditions were instructed on the development of critical thinking questions. Therefore, to ensure that all participants had sufficient domain knowledge for situation model production, supplementary text was necessitated.

Following the "Blood" text, all participants received the experimental text entitled "Heart Disease". Participants either received the high coherence (CM) text or the low coherence (cm) text, depending upon their predetermined treatment assignment. Participants were not informed that there were two forms of the text. They were instructed to work individually and were given exactly 5 minutes to read the experimental text.

After reading the "Heart Disease" text, participants were given their strategy prompt cards and asked to start individually generating questions. They were instructed to think of questions using the strategies for which they were trained. All participants were given question/response sheets on which they are to record their questions and the responses to those questions

(Appendix O). It was emphasized that the text entitled "Heart Disease" is of primary concern, but that questions may *link* to previously read text(s) (i.e., the "Blood" text or any other prior domain knowledge). Participants were allowed to refer back to both the preliminary and experimental texts as much as they so chose. All participants were given exactly 20 minutes to complete this task. Participants in the generic question stem condition and signal word condition were reminded that they must use all 6 of the question stems/signal words and that they must not use the same question stem/signal word more than once. Once the question sheets and responses were completed, they were collected by the researcher. This prevented the participants from studying their formulated questions/responses prior to testing.

The calibration of comprehension questionnaire was then administered. Participants were told to read each question and circle a single number on the accompanying scale to report how confident they were that they could correctly answer the question being asked. Prediction questions were matched with questions on the post-test. Participants were given exactly 5 minutes. It was followed by the summary recall test which asked the participants to summarize what they had read in the text entitled "Heart Disease" using an essay format. Thus, they were being asked to state the important information in the text in a briefer form. Participants were given exactly 10 minutes.

The post-test was then administered. It asked the participants to answer all of the questions carefully and fully. They were allowed to use point form so long as their answers were clear and concise. Exactly 30 minutes was allotted for this task. Finally, the post-reading sorting task was given. Because it was identical to the pre-reading sorting task, participants were simply reminded of the procedures. Again, exactly 10 minutes was given for completion.

Analysis of Data

Three-factor multiple analysis of variance (MANOVA) was used to identify which treatment group (i.e., generic question stems, signal words, unguided questioning), domain knowledge level (i.e., high domain knowledge, low domain knowledge) and/or level of text coherence (i.e., high text coherence, low text coherence) affects a reader's memory and ability to learn from text. Where significant multivariate main effects and interactions were found, univariate analysis of variance (ANOVA) was used to identify the dependent variable(s) contributing to the significant multivariate result. An analysis of a single degree of freedom was used to identify the specific simple effects (i.e., pair-wise treatment differences) and 2 x 2 interactions. Finally, Tukey's post hoc procedure was used to determine those cells where significant differences occurred. For all hypotheses, statistical tests were set at an alpha probability of .05. This ensures reasonable assurance against Type I errors as well as consistency in producing power statistics for all the tests.

Scoring was done by the researcher. Scoring keys for each instrument were designed prior to testing and strictly followed. Because the posttest and summary recall require some subjective judgments, a second rater was employed. Therefore, reliability in scoring was assessed on the summaries and post tests of 24 participants who represented two randomly selected participants from each cell of the 3 x 2 x 2 design (treatment, text coherency, and domain knowledge). Interrater reliability ranged from .81 to .97. For detailed description of reliability coefficients see Tables 2 and 3.

Table 2

Inter-Rater Reliability Coefficients* for Posttest

Posttest Total Score	Posttest Bridging Inference	Posttest Elaborative Inference	Posttest Textbase	Posttest Problem Solving
.88	.81	.83	.86	.85

*Correlations are significant at the .05 level (2-tailed).

Table 3

Inter-Rater Reliability Coefficients* for Summary Recall

Propositions	Generalizations	Elaborations	Reorderings	Level 1	Level 2	Level 3	Level 4
.87	.87	.96	.84	.97	.92	.88	.86

*Correlations are significant at the .05 level (2-tailed).

CHAPTER FOUR

RESULTS

This study was guided by three questions. Firstly, can we expect high or low domain knowledge participants who are reading high coherence or low coherence expository text to attain higher measurable outcomes on memory (i.e., textbase model) and learning (i.e., situation model) variables when using generic question stems as compared to signal words and/or unguided questioning techniques? Secondly, will there be interaction effects on these variables suggesting that combinations of domain knowledge (i.e., high and low) and text coherency (i.e., high coherence and low coherence) are uniquely affected by generic question stems, signal words, or unguided questioning? And finally, will the use of generic questions stems lead to superior calibration of comprehension scores as compared to those generated by signal words and/or unguided questioning conditions?

In this chapter, outcomes are reported according to the data collection sequence, the guiding research questions, and the research instruments employed. Results from the multivariate, univariate, single degrees of freedom, Tukey's post hoc, and gamma/Pearson correlations analyses are reported as necessary.

Research Questions #1 and #2

The following measures provide results pertaining to the first and second

questions guiding this study. Statistical procedures and outcomes are reported and briefly discussed. An in-depth discussion of the results as they pertain to each specific research question can be found in Chapter 5.

Sorting Data. A repeated measures analysis was conducted using the pre- and post-treatment sorting data as a within-subjects factor and treatment, domain knowledge, and text coherence as between-subjects factors. Combined means and standard deviations for the pre- and post-treatment sorting data for the cells are reported in Table 4. Combined means and standard deviations for the pre- and post-treatment sorting data for Treatment, Domain Knowledge, and Text Coherence are reported in Table 5.

Results of the repeated measures analysis show significant main effects for sorting data [$F(1, 51) = 10.46, p < .05$, effect size = .170, power = .89] and interaction effects for sorting data by treatment [$F(2, 51) = 5.21, p < .05$, effect size = .170, power = .81]. As evidenced by the changes in sorting scores alone, participants significantly improved from pre- ($M = .27$) to post-treatment ($M = .36$). No significant differences were found for the between-subjects factors.

To determine where significant multivariate interactions occurred for the sorting data by treatment factor, a single degree of freedom analysis was conducted. Two analyses were run. For the two parameters, Treatment #1 versus #3 and Treatment #3 and #2, the multivariate effects were significant ($p < .05$) [$F(1, 51) = 10.30$ and $F(1, 51) = 8.69$, effect size = .170, power = .81, respectively]. Their power indicates a reasonable robustness of the tests to detect Type II errors.

Table 4

Means and Standard Deviations By Treatment, Domain Knowledge, Text Coherence, and Pre- and Post-Treatment Sorting Data

Treatment	Domain knowledge	Text coherence	n	Pre-Treatment Sort	Post-Treatment Sort
1	high	high	6	.32 (.15)	.34 (.11)
		low	6	.28 (.08)	.30 (.13)
	low	high	4	.28 (.10)	.27 (.17)
		low	5	.30 (.27)	.24 (.18)
2	high	high	5	.25 (.21)	.27 (.16)
		low	5	.21 (.10)	.39 (.19)
	low	high	6	.29 (.12)	.35 (.14)
		low	3	.27 (.06)	.28 (.08)
3	high	high	6	.22 (.06)	.49 (.25)
		low	6	.31 (.17)	.51 (.16)
	low	high	5	.23 (.09)	.37 (.19)
		low	6	.29 (.08)	.43 (.14)

Note. Treatment 1 = generic question stems; Treatment 2 = signal word; Treatment 3 = unguided questioning. Values enclosed in parentheses represent standard deviations.

Table 5

Combined Means and Standard Deviations for Treatment, Domain Knowledge, and Text Coherence: Independent Variables (Factors)

<u>Variable</u>	<u>Treatment</u>			<u>Domain Knowledge</u>		<u>Text Coherence</u>	
	1	2	3	1	2	1	2
Pre-Treatment Sort	.30 (.16)	.25 (.13)	.27 (.11)	.27 (.13)	.28 (.13)	.26 (.12)	.28 (.14)
Post-Treatment Sort	.29 (.14)	.33 (.15)	.45 (.18)	.39 (.18)	.33 (.16)	.36 (.18)	.37 (.17)

Note. Treatment
 1 = generic question stems
 2 = signal word
 3 = unguided questioning

Domain Knowledge
 1 = high domain knowledge
 2 = low domain knowledge

Text Coherence
 1 = high text coherence
 2 = low text coherence

Values enclosed in parentheses represent standard deviations.

Therefore, it appears that participants in the signal word and unguided questioning conditions improved in their sorting scores from pre-treatment to post-treatment while this was not true of the generic question stem participants. Moreover, the unguided questioning participants performed significantly better than participants in either the generic question stem or signal word conditions. Why the structured questioning techniques caused participants to digress from the sorting matrix is puzzling. Perhaps they incited participants to activate connections which are either too general or too remote. From an observational perspective, participants in the generic question stem condition frequently created a greater number of categories which made distant connections between terms. Category labels included "chemical compounds", "causes of heart attacks or heart diseases", "diseases of the heart", "terms related to the heart", "disease or defect related words", "lung related words", "terms related to rheumatic fever", and so forth. Comparatively, the unguided questioning participants typically made specific connections which reflect the sorting matrix. They were labeled as "parts of the heart", "internal organs/other body parts", "tumors", "treatments for heart disease", and so forth. What ever the reasons, the less structured the questioning technique, the more accurate the sorting score.

Post-test. Analysis of variance (ANOVA) was used to determine treatment effects on the Post-test total score while multivariate analysis of variance (MANOVA) was conducted using question types (i.e., bridging inference, elaborative inference, textbase, and problem solving). Combined means and standard deviations for the Post-test and question types for the cells are reported in Table 6. Combined means and standard deviations for the Post-

Table 6

Means and Standard Deviations By Treatment, Domain Knowledge, Text Coherence, and Question Type

Treatment	Domain Knowledge	Text Coherence	n	Post-test Total Score	Textbase	Elaborative Inference	Bridging Inference	Problem Solving
1	high	high	6	52.7 (18.5)	13.7 (5.6)	12.2 (4.8)	13.0 (4.9)	14.0 (4.9)
		low	6	61.5 (10.8)	16.4 (3.5)	15.4 (2.7)	14.2 (2.5)	15.6 (3.7)
	low	high	4	36.2 (18.1)	8.0 (6.7)	8.1 (4.7)	9.5 (4.1)	11.2 (3.8)
		low	5	45.4 (8.3)	13.3 (2.2)	11.1 (2.2)	11.8 (3.3)	9.6 (3.2)
2	high	high	5	51.1 (8.9)	13.2 (3.3)	12.4 (4.1)	11.8 (2.9)	13.7 (1.5)
		low	5	51.3 (8.0)	13.6 (4.7)	12.4 (2.6)	12.4 (2.6)	12.6 (2.8)
	low	high	6	34.0 (10.1)	8.3 (3.9)	7.6 (3.7)	9.4 (2.3)	9.6 (3.5)
		low	3	32.0 (5.6)	12.2 (2.6)	5.1 (1.1)	9.7 (1.5)	5.8 (3.2)
3	high	high	6	58.2 (11.6)	15.2 (3.1)	13.0 (3.5)	15.0 (3.3)	15.6 (4.1)
		low	6	43.9 (11.7)	10.0 (4.4)	10.9 (2.9)	10.2 (2.7)	12.5 (3.6)
	low	high	5	43.6 (8.8)	13.7 (1.9)	9.4 (1.9)	11.1 (2.5)	10.0 (4.0)
		low	6	38.3 (5.4)	10.0 (3.7)	7.3 (1.9)	11.2 (.7)	10.8 (4.0)

Note. Treatment 1 = generic question stems; Treatment 2 = signal word; Treatment 3 = unguided questioning. Values enclosed in parentheses represent standard deviations.

test and question types for Treatment, Text Coherence and Domain Knowledge are reported in Table 7. Correlation coefficients for question types are reported in Appendix P.

Results of the Factorial ANOVA concerning the Post-test total score show a significant main effect for domain knowledge [$F(1, 51) = 25.99, p < .05$, effect size = .338, power = .99] and a significant two-way interaction between treatment and text coherence [$F(1, 51) = 3.68, p < .05$, effect size = .126, power = .65]. No significant differences were found for the treatment or text coherence main effects. However, the mean of the generic question stem condition ($M = 50.4$), although not significantly different, was higher than the signal word condition ($M = 42.7$) and unguided questioning condition ($M = 46.1$).

With respect to domain knowledge differences, high domain knowledge participants had a significantly higher mean in Post-test total scores ($M = 53.2$) than their low domain knowledge counterparts ($M = 38.6$). To ascertain where the significant univariate interactions occurred, a single degree of freedom analysis was conducted. In order to obtain results for all three pairs of the treatment factor with text coherence, two analyses were run. For the two interaction parameters, Treatment #1 versus #3 and text coherence and Treatment #3 versus #2 and text coherence, the univariate effects were significant ($p < .05$) [$F(1,51) = 7.35$, effect size = .126, power = .65 and $F(1,62) = 5.28$, effect size = .126, power = .65], respectively.

For the purposes of analyzing meaningful univariate pairwise differences of the treatment and text coherence combined cell means (see Table 8), the Tukey method of multiple comparisons was employed. The mean of the generic question stem condition/low coherence text was significantly

Table 7

Combined Means and Standard Deviations for Treatment, Domain Knowledge, and Text Coherence: Independent Variables (Factors)

<u>Variable</u>	<u>Treatment</u>			<u>Domain Knowledge</u>		<u>Text Coherence</u>	
	1	2	3	1	2	1	2
Post-Test Total Score	50.4 (16.2)	42.7 (12.2)	46.1 (11.9)	53.2 (12.7)	38.6 (10.2)	46.5 (15.1)	46.5 (12.3)
Textbase	13.3 (5.2)	11.6 (4.2)	12.2 (3.9)	13.7 (4.4)	10.8 (4.1)	12.2 (4.8)	12.6 (4.2)
Elaborative Inference	12.1 (4.2)	9.7 (4.2)	10.2 (3.3)	12.7 (3.5)	8.3 (3.1)	10.6 (4.1)	10.8 (3.9)
Bridging Inference	12.4 (3.9)	10.9 (2.7)	11.9 (3.0)	12.8 (3.5)	10.5 (2.5)	11.8 (3.8)	11.7 (2.7)
Problem Solving	12.9 (4.4)	10.9 (3.8)	12.3 (4.2)	14.0 (3.6)	9.8 (3.6)	12.4 (4.2)	11.7 (4.2)
<u>Note.</u>	<u>Treatment</u>			<u>Domain Knowledge</u>		<u>Text Coherence</u>	
	1 = generic question stems			1 = high domain knowledge		1 = high text coherence	
	2 = signal word			2 = low domain knowledge		2 = low text coherence	
	3 = unguided questioning						

Values enclosed in parentheses represent standard deviations.

Table 8

**Tukey's Pairwise Comparisons for the Treatment x Text Coherence Combined
Cell Means for the Post-Test Total Score**

Means of Post-Test Total Score								
Treatment	Text Coherence	Mean	high (1) 46.1	low (1) 54.2	high(2) 41.8	low (2) 44.1	high(3) 51.5	low(3) 41.1
<hr/>								
1	high	46.1	-	-	-	-	-	-
	low	54.2	3.3	-	-	-	-	-
2	high	41.8	1.7	4.9	-	-	-	-
	low	44.1	.8	3.9	.9	-	-	-
3	high	51.5	2.2	1.1	3.9	3.0	-	-
	low	41.1	2.1	5.4*	.2	1.2	4.3*	-

* Indicate the observed q is significant at .05

higher than the mean of the unguided questioning condition/low coherence text. Similarly, the mean of the unguided questioning condition/high coherence text was significantly higher than the mean of the unguided questioning condition/low coherence text. On the basis of this information, it is evident that high coherence text is necessary for knowledge acquisition where there is unguided questioning, but the utilization of generic question stems is required when reading low coherence text.

Results of the 3 x 2 x 2 MANOVA with four dependent variables (bridging inference, elaborative inference, textbase, and problem solving) show a significant main effect for domain knowledge [$F(4, 48) = 8.97, p < .05$, effect size = .428, power = .99] and a significant two-way interaction between treatment and text coherence [$F(8, 98) = 2.51, p < .05$, effect size = .176, power = .89]. With respect to domain knowledge, univariate main effects were significant ($p < .05$) for all four question types (i.e., textbase [$F(1, 51) = 7.27$, effect size = .125, power = .75]; elaborative inference [$F(1, 51) = 30.83$, effect size = .377, power = 1.0]; bridging inference [$F(1, 51) = 8.95$, effect size = .149, power = .83]; and problem solving [$F(1, 51) = 22.37$, effect size = .305, power = .99]). As can be seen in Table 7, high domain knowledge participants significantly outperformed low domain knowledge participants. Univariate interactions were only significant ($p < .05$) for textbase [$F(2, 51) = 6.65$, effect size = .207, power = .90] and elaborative inference questions [$F(2, 51) = 4.01$, effect size = .136, power = .69]. No significant differences were found for the treatment or text coherence main effects. However, the means of the generic question stem condition, although not significantly different, were higher than the other treatment conditions on all question types (see Table 7).

To ascertain where significant multivariate interactions occurred between

the terms treatment and text coherence, a single degree of freedom analysis was conducted. Again, two analyses were run. For the two interaction parameters, Treatment #1 versus #3 and text coherence and Treatment #3 versus #2 and text coherence, the multivariate effects were significant ($p < .05$) [$F(4, 48) = 3.64$, effect size = .233, power = .84 and $F(4, 48) = 3.73$, effect size = .237, power = .85]. Their power indicates a reasonable robustness of the tests to detect Type II errors.

Univariate single degree of freedom partition results for the interaction are shown in Table 9. The parameters for textbase scores were found to be reliable ($p < .05$) for two of the possible combinations of the treatment levels with the text coherence term. The parameter for elaborative inference scores was found to be reliable ($p < .05$) for one combination (see Table 9).

The Tukey method of multiple comparisons was used to analyze meaningful univariate pairwise differences in the combined means for treatment and text coherence interaction. With respect to the textbase question type, the mean of the generic question stem condition/low coherence text was significantly higher than the mean of the unguided questioning condition/low coherence text. Similarly, the mean of the unguided questioning condition/high coherence text was significantly higher than the mean of the unguided questioning/low coherence text as well as the signal word condition/high coherence text. In regards to elaborative inferencing question responses, the mean of the generic question stem condition/low coherence text was significantly higher than the mean of the unguided questioning condition/low coherence text as well as the generic question stem condition/high coherence text (see Table 10). Therefore, the generic question stem condition /low coherence text outperformed all other cells with some

Table 9

Treatment x Text Coherence Interaction: Significant Simple Effects for Univariate Analysis of Variance

Parameter	Variable	F (1, 51)	Power
1st			
(Treatment #1 - Treatment #3)	Textbase	12.20*	.93
	Elaborative Inference	7.07*	.74
2nd			
(Treatment #3 - Treatment #2)	Textbase	12.54*	.93

* $p < .05$

Table 10

**Tukey's Pairwise Comparisons for the Treatment x Text Coherence Combined
Cell Means for the Textbase and Elaborative Inference Scores**

Means of Textbase Scores								
Treatment	Text Coherence	Mean	high (1) 11.4	low (1) 15.0	high(2) 10.5	low (2) 13.1	high(3) 14.5	low(3) 10.0
1	high	11.4	-	-	-	-	-	-
	low	15.0	4.1	-	-	-	-	-
2	high	10.5	1.0	4.9	-	-	-	-
	low	13.1	1.8	2.1	2.9	-	-	-
3	high	14.5	3.6	.5	4.5*	1.7	-	-
	low	10.0	1.7	5.8*	.6	3.5	5.2*	-

* Indicate the observed q is significant at .05

Means of Elaborative Inferencing Scores								
Treatment	Text Coherence	Mean	high (1) 10.5	low (1) 13.5	high(2) 9.8	low (2) 9.7	high(3) 11.4	low(3) 9.1
1	high	10.5	-	-	-	-	-	-
	low	13.5	4.3*	-	-	-	-	-
2	high	9.8	1.0	5.1	-	-	-	-
	low	9.7	1.2	5.3	.1	-	-	-
3	high	11.4	1.2	3.0	2.3	2.4	-	-
	low	9.1	2.1	6.4*	.9	.8	3.3	-

* Indicate that the observed q is significant at .05

significant differences. Thus, it appears that the combination of highly structured questioning and low coherence text induces a deeper level of processing.

Summary recall.

(i). Inference analysis. Analysis of variance (ANOVA) was used to determine treatment effects on the total number of propositions cited while multivariate analysis of variance (MANOVA) was used to determine the total number of generalizations, elaborations, and reorderings. Combined means and standard deviations for the propositions and generalizations, elaborations, and reorderings for the cells are reported in Table 11. Combined means and standard deviations for the propositions and generalizations, elaborations, and reorderings for Treatment, Domain Knowledge, and Text Coherence are reported in Table 12.

Results of the Factorial ANOVA concerning the total number of propositions cited show significant main effects for domain knowledge [$F(1, 51) = 9.51$, $p < .05$, effect size = .157, power = .86] and treatment [$F(2, 51) = 3.60$, $p < .05$, effect size = .123, power = .64].

With respect to domain knowledge differences, high domain knowledge participants produced more propositions ($M = 11.5$) than their low domain knowledge counterparts ($M = 8.1$). For treatment effects, participants in the unguided questioning condition generated the greatest number of propositions ($M = 11.9$) followed by the generic questioning stem ($M = 9.3$) and signal word conditions ($M = 8.4$).

To ascertain where the significant univariate pairwise treatment differences occurred, a single degree of freedom analysis was conducted. In

Table 11

Means and Standard Deviations By Treatment, Domain Knowledge, Text Coherence, and Summary Recall Data

Treatment	Domain knowledge	Text coherence	n	Propositions	Generalizations	Elaborations	Reorderings
1	high	high	6	10.4 (6.2)	4.6 (2.3)	1.2 (1.1)	1.7 (.9)
		low	6	10.8 (4.6)	7.9 (3.8)	3.2 (3.3)	2.5 (1.8)
	low	high	4	6.5 (6.5)	7.1 (1.9)	.5 (.6)	1.4 (.5)
		low	5	8.2 (6.4)	5.8 (4.9)	2.3 (1.5)	3.3 (2.4)
2	high	high	5	10.5 (2.9)	4.3 (1.8)	.2 (.4)	2.7 (.9)
		low	5	9.9 (3.4)	4.7 (1.1)	.0 (.0)	1.4 (1.1)
	low	high	6	7.7 (2.0)	4.6 (3.4)	.6 (.7)	.2 (.6)
		low	3	3.7 (.6)	3.8 (.3)	1.0 (1.0)	2.7 (3.8)
3	high	high	6	13.6 (6.0)	5.0 (1.3)	1.7 (1.4)	.6 (.7)
		low	6	13.5 (5.5)	4.6 (1.5)	1.1 (1.6)	3.4 (1.7)
	low	high	5	8.9 (5.8)	5.0 (4.0)	.5 (.7)	.8 (1.0)
		low	6	11.1 (.8)	4.2 (2.6)	.3 (.5)	.9 (.9)

Note. Treatment 1 = generic question stems

Treatment 2 = signal word

Treatment 3 = unguided questioning

Values enclosed in parentheses represent standard deviations.

Table 12

Combined Means and Standard Deviations for Treatment, Domain Knowledge, and Text Coherence: Independent Variables (Factors)

<u>Variable</u>	<u>Treatment</u>			<u>Domain Knowledge</u>		<u>Text Coherence</u>	
	1	2	3	1	2	1	2
Propositions	9.3 (5.7)	8.4 (3.3)	11.9 (5.0)	11.5 (4.9)	8.1 (4.5)	9.8 (5.2)	10.1 (4.8)
Generalizations	6.3 (3.5)	4.4 (2.1)	4.7 (2.3)	5.2 (2.4)	5.1 (3.2)	5.0 (2.6)	5.3 (3.1)
Elaborations	1.9 (2.1)	.4 (.6)	.9 (1.2)	1.3 (1.9)	.8 (1.0)	.8 (.9)	1.4 (2.0)
Reorderings	2.3 (1.7)	1.6 (1.8)	1.5 (1.6)	2.1 (1.5)	1.4 (1.8)	1.2 (1.1)	2.3 (2.0)

<u>Note.</u>	Treatment	Domain Knowledge	Text Coherence
	1 = generic question stems	1 = high domain knowledge	1 = high text coherence
	2 = signal word	2 = low domain knowledge	2 = low text coherence
	3 = unguided questioning		

Values enclosed in parentheses represent standard deviations.

order to obtain results for all three treatment factors, two analyses were run. For one parameter, Treatment #3 and #2, the univariate effect was significant ($p < .05$) [$F(1, 51) = 6.85$, effect size = .123, power = .64]. Therefore, participants in the unguided questioning condition produced on average a significantly higher number of propositions than the participants in the signal word condition.

A 3 x 2 x 2 MANOVA with three dependent variables was conducted to determine the effects of treatment, domain knowledge, and text coherence on generalizations, elaborations, and reorderings. Results indicate significant main effects for treatment [$F(6, 96) = 2.60$, $p < .05$, effect size = .139, power = .83] and text coherence [$F(3, 49) = 3.30$, $p < .05$, effect size = .167, power = .72]. A significant three-way interaction was also detected between treatment, text coherence, and domain knowledge [$F(6, 96) = 2.51$, $p < .05$, effect size = .176, power = .89]. With respect to treatment, univariate main effects were significant ($p < .05$) for elaborations [$F(2, 51) = 4.90$, effect size = .161, power = .78] while reorderings were significant for text coherence [$F(1, 51) = 8.80$, effect size = .147, power = .83] and the three way interaction of treatment, text coherence, and domain knowledge [$F(2, 51) = 6.20$, effect size = .195, power = .87]. No significant differences were detected for the domain knowledge main effects.

In regards to text coherence, participants reading low coherence text generated significantly more reorderings ($M = 2.3$) than those generated by their high coherence counterparts ($M = 1.2$). To ascertain where the significant univariate interactions occurred for treatment main effects and treatment, domain knowledge, and text coherence, a single degree of freedom analysis was conducted. In order to obtain results for all three pairs of the treatment factor with the elaboration term, two analyses were run. For the two parameters

of the treatment levels (i.e., Treatment #1 and #3 and Treatment #1 and #2), the univariate effects were significant ($p < .05$) [$F(3, 49) = 3.06$, effect size = .158, power = .68 and $F(3, 49) = 4.33$, effect size = .210, power = .84], respectively.

Single degree of freedom analyses were also run for the treatment, domain knowledge, and text coherence interaction. Only one interaction parameter, Treatment #3 and #2 and domain knowledge and text coherence was significant ($p < .05$) [$F(3, 49) = 3.89$, effect size = .192, power = .79].

Univariate single degree of freedom partition results for the treatment main effects and the interaction are shown in Table 13. The parameters for elaboration scores were found to be reliable ($p < .05$) for two of the possible combinations of the treatment levels. The parameter for reordering scores was found to be reliable ($p < .05$) for only one combination of the treatment with the text coherence and domain knowledge terms (see Table 13).

For the purposes of analyzing meaningful univariate pairwise differences, the Tukey method of multiple comparisons was employed. The reordering mean of the high domain knowledge/low text coherence participants of Treatment #3 was significantly higher than the mean of the low domain knowledge/high text coherence participants of Treatment #2, the high domain knowledge/high text coherence participants of Treatment #3, the low domain knowledge/high text coherence participants of Treatment #3, the low domain knowledge/low text coherence participants of Treatment #3, and the high domain knowledge/low text coherence participants of Treatment #2. The mean of the high domain knowledge/high text coherence participants of Treatment #2 was also significantly higher than the mean of the low domain knowledge/high text coherence participants of Treatment #2, the high domain knowledge/high text coherence participants of Treatment #3, the low domain knowledge and low

Table 13

Treatment and Treatment, Text Coherence x Domain Knowledge Interaction:
Significant Simple Effects for Univariate Analysis of Variance

Parameter	Variable	F (1, 51)	Power
1st (Treatment #1 - Treatment #3)	Elaborations	4.43*	.54
2nd (Treatment #1 - Treatment #2)	Elaborations	9.36*	.85
2nd (Treatment #3 - Treatment #2)	Propositions	6.85*	.64
	Reorderings	10.94*	.89

* $p < .05$

text coherence participants of Treatment #3, and the low domain knowledge/high text coherence participants of Treatment #3. Finally, the mean of the low domain knowledge/low text coherence participants of Treatment #2 was significantly higher than the mean of the low domain knowledge/high text coherence participants of Treatment #2, the high domain knowledge/high text coherence participants of Treatment #3, and the low domain knowledge/high text coherence participants of Treatment #3 (See Table 14).

Therefore, it appears that propositions are more easily produced from memory by those with high domain knowledge and by those within the unguided questioning condition. This stands to reason as those with more domain knowledge can recall specific information more easily as it matches already existing schemata. With respect to the superiority of the unguided questioning condition, this too follows as participants were not required to make connections within the text and between other knowledge sources thereby inciting them to encode detail.

In regards to elaborations, generic question stem participants generated more elaborations than either of the other two conditions. This too follows as the question stems require participants to link information within the text to additional information sources. This is not the case with signal word and unguided questioning strategies.

Finally, with respect to reorderings, those reading low coherence text had significantly more reorderings than those reading high coherence text. This can be expected as low coherence text lacks strong macro- and micro-structures. Thus, the reader is left to encode the information in his or her memory in potentially different sequences than those found within the text. Also, as found within the literature (McNamara et al., 1996), high domain knowledge/low

Table 14

Tukey's Pairwise Comparisons for the Treatment Combined Cell Means for the Elaborations and Reorderings

Means of Elaborations				
Treatment	Mean	1	2	3
		1.8	.4	.9
1	1.8	-	-	-
2	.4	4.4*	-	-
3	.9	3.1*	1.5	-

* p < .05

Means of Reorderings															
Treat.	Domain knowl.	Text coher.	Mean	H/H (1)	H/L (1)	L/H (1)	L/L (1)	H/H (2)	H/L (2)	L/H (2)	L/L (2)	H/H (3)	H/L (3)	L/H (3)	L/L (3)
				1.7	2.5	1.4	3.3	2.7	1.4	.2	2.7	.6	3.4	.8	.92
1	high	high	1.7	-	-	-	-	-	-	-	-	-	-	-	-
		low	2.5	2.3	-	-	-	-	-	-	-	-	-	-	-
	low	high	1.4	1.1	3.4	-	-	-	-	-	-	-	-	-	-
		low	3.3	4.7	2.4	5.8	-	-	-	-	-	-	-	-	-
2	high	high	2.7	2.9	.6	1.3	1.8	-	-	-	-	-	-	-	-
		low	1.4	1.1	3.3	.1	5.7	3.9	-	-	-	-	-	-	-
	low	high	.2	4.5	6.8	3.4	9.2	7.4*	3.4	-	-	-	-	-	-
		low	2.7	2.8	.5	3.9	1.9	.1	3.8	7.3*	-	-	-	-	-
3	high	high	.6	3.7	6.1	2.5	8.6	6.6*	2.5	1.0	6.4*	-	-	-	-
		low	3.4	5.3	2.9	6.5	.4	2.2	6.2*	9.8*	2.3	8.8*	-	-	-
	low	high	.8	3.0	5.4	1.8	7.9	5.9*	1.8	1.7	5.8*	.7	8.1*	-	-
		low	.9	2.6	5.0	1.4	7.6	5.5*	1.5	2.1	5.4*	1.0	7.7*	.4	-

p < .05

coherence text participants are forced to process information more deeply than any other combination of domain knowledge/text coherence. This is because incoherency within text structure forces them to engage in more inferential processing. Moreover, it appears that those high domain knowledge/low coherence text participants who did not receive additional guidance in their questioning tactics (i.e., unguided questioning condition) reordered the material significantly more often than most other treatment combinations. However, when given a questioning strategy such as signal words, those with low domain knowledge/low text coherence were able to outperform their unguided questioning/high domain knowledge/high text coherence counterparts. This suggests that the provision of a structured questioning strategy is successful in helping those entering a learning situation with few advantages versus those with the greatest number of advantages.

(ii). Levels analysis. Multivariate analysis of variance (MANOVA) was used to determine treatment effects on Levels 1, 2, 3, and 4. Combined means and standard deviations for each level for the cells and Treatment are reported in Table 15. Combined means and standard deviations for each level and Treatment, Domain Knowledge, and Text Coherence are reported in Table 16.

Results of 3 x 2 x 2 MANOVA with four dependent variables (Level 1, Level 2, Level 3, and Level 4) show significant main effects for domain knowledge [$F(4, 48) = 3.75, p < .05$, effect size = .238, power = .86] and text coherence [$F(4, 48) = 5.78, p < .05$, effect size = .325, power = .97]. A significant three way interaction was also detected between treatment, domain knowledge, and text coherence [$F(8, 94) = 2.11, p < .05$, effect size = .150, power = .81] .

Table 15

Means and Standard Deviations By Treatment, Domain Knowledge, Text Coherence, and Levels Analysis Data

Treatment	Domain knowledge	Text coherence	n	Level 1	Level 2	Level 3	Level 4
1	high	high	6	41.7 (49.1)	38.9 (44.3)	47.7 (19.9)	18.6 (6.3)
		low	6	37.5 (20.9)	11.1 (17.2)	36.4 (20.1)	31.6 (12.8)
	low	high	4	93.8 (12.5)	58.3 (41.9)	40.9 (28.5)	15.7 (10.3)
		low	5	30.0 (27.4)	20.0 (29.8)	20.9 (13.5)	20.5 (10.7)
2	high	high	5	65.0 (48.7)	66.7 (47.1)	27.3 (22.3)	21.3 (6.4)
		low	5	40.0 (22.3)	19.9 (18.2)	39.1 (20.7)	23.3 (3.9)
	low	high	6	25.0 (27.4)	38.9 (13.6)	24.9 (13.4)	19.9 (5.6)
		low	3	75.0 (43.3)	22.2 (38.5)	15.1 (13.9)	7.3 (1.9)
3	high	high	6	45.8 (51.0)	94.4 (13.6)	57.6 (15.9)	26.9 (21.1)
		low	6	33.3 (25.8)	22.2 (17.2)	27.3 (22.8)	29.7 (12.9)
	low	high	5	70.0 (27.4)	59.9 (27.9)	36.4 (21.1)	15.1 (9.6)
		low	6	25.0 (27.4)	41.7 (29.3)	25.7 (18.8)	23.5 (3.3)

Note. Treatment 1 = generic question stems

Treatment 2 = signal word

Treatment 3 = unguided questioning

Values enclosed in parentheses represent standard deviations.

Table 16

Combined Means and Standard Deviations for Treatment, Domain Knowledge, and Text Coherence: Independent Variables (Factors)

<u>Variable</u> <u>Coherence</u>	<u>Treatment</u>			<u>Domain Knowledge</u>		<u>Text</u>	
	1	2	3	1	2	1	2
Level 1	47.6 (37.8)	47.4 (38.1)	42.4 (36.5)	43.4 (37.1)	48.3 (37.1)	53.9 (42.6)	37.1 (28.0)
Level 2	30.2 (36.4)	38.6 (33.8)	54.3 (34.9)	42.2 (40.4)	40.8 (30.7)	59.4 (36.6)	23.1 (24.6)
Level 3	36.8 (21.5)	27.7 (18.5)	36.8 (22.6)	39.6 (21.8)	27.6 (18.7)	39.5 (21.8)	28.4 (19.2)
Level 4	22.2 (11.4)	19.2 (7.2)	24.2 (13.7)	25.4 (12.4)	18.0 (8.5)	19.9 (11.2)	24.2 (11.1)

Note. Treatment
1 = generic question stems
2 = signal word
3 = unguided questioning

Domain Knowledge
1 = high domain knowledge
2 = low domain knowledge

Text Coherence
1 = high text coherence
2 = low text coherence

Values enclosed in parentheses represent standard deviations.

With respect to univariate main effects, domain knowledge was significant ($p < .05$) with Levels 3 [$F(1,51) = 5.55$, effect size = .098, power = .64] and 4 [$F(1,51) = 9.23$, effect size = .153, power = .85] while Levels 2 [$F(1,51) = 23.50$, effect size = .315, power = .99] and 3 [$F(1,51) = 5.41$, effect size = .096, power = .63] were significant with the text coherence factor ($p < .05$). For the three way interaction between treatment, text coherence, and domain knowledge, there was a significant univariate effect for Level 1 [$F(2,51) = 5.10$, effect size = .166, power = .80].

In regards to domain knowledge, high domain knowledge participants generated significantly more Level 3 and 4 responses than their low domain counterparts (see Table 16). Moreover, those participants reading high text coherence generated significantly more Level 2 and 3 responses than their low text coherence counterparts (see Table 16).

To ascertain where the significant univariate interactions occurred for the Level 1 responses of treatment, domain knowledge, and text coherence, a single degree of freedom analysis was conducted. In order to obtain results for all three pairs of the treatment factor, two analyses were run. Only one interaction parameter, Treatment #2 and #3 and domain knowledge and text coherence was significant ($p < .05$) [$F(4, 48) = 2.97$, effect size = .199, power = .75].

For the purposes of analyzing meaningful univariate pairwise differences, the Tukey method of multiple comparisons indicates that the mean of the low domain knowledge/low text coherence participants of Treatment #2 was significantly higher than the mean of the low domain knowledge/high text coherence participants of Treatment #2, the low domain knowledge/low text coherence participants of Treatment #3, and the high domain knowledge/low

text coherence participants of Treatment #3 and Treatment #2. Similarly, the mean of the low domain knowledge/high text coherence participants of Treatment #3 was significantly higher than the mean of the low domain knowledge/high text coherence participants of Treatment #2 and the low domain knowledge/low text coherence participants of Treatment #3. Finally, the mean of the high domain knowledge/high text coherence members of Treatment #2 was significantly higher than the mean of the low domain knowledge/high text coherence participants of Treatment #2 and the low domain knowledge/low text coherence participants of Treatment #3 (see Table 17).

Therefore, it appears that high domain knowledge is an asset when recalling inferred or explicit subheadings (i.e., Level 3) and details (i.e., Level 4) from the text. In regards to recollections of details this is certainly not surprising as it is easier to encode and retrieve data which corresponds with previously existing schemata. With respect to text coherence, those reading highly coherent micro- and macro-structure text recalled more Level 2 and Level 3 data. This was to be expected as the high coherence text explicitly stated the three major subtopics (i.e., Level 2) and was more explicit when stating and/or outlining subheadings (i.e., Level 3). Finally, the presence of a semi-structured questioning strategy assisted the low domain knowledge/low text coherence participants to state significantly more often what the text was about (i.e., Level 1).

Research Question #3

The following measure provides results pertaining to the third question

Table 17

Tukey's Pairwise Comparisons for the Treatment x Text Coherence x Domain Knowledge Combined Cell Means for the Level 1 Means of Treatments #2 and #3

				Means of Level 1							
				H/H (2)	H/L (2)	L/H (2)	L/L (2)	H/H (3)	H/L (3)	L/H (3)	L/L (3)
Treat.	Domain knowl.	Text coher.	Mean	65.0	40.0	25.0	75.0	45.8	33.3	70.0	25.0
2	high	high	65.0	-	-	-	-	-	-	-	-
		low	40.0	3.3	-	-	-	-	-	-	-
	low	high	25.0	5.3*	1.9	-	-	-	-	-	-
		low	75.0	1.3	4.7*	6.6*	-	-	-	-	-
3	high	high	45.8	2.5	.8	2.8	3.9	-	-	-	-
		low	33.3	4.2	1.7	1.1	5.5*	1.7	-	-	-
	low	high	70.0	1.3	3.9	5.9*	.7	3.2	5.5*	-	-
		low	25.0	5.3*	1.9	.0	6.6*	2.8	1.1	6.6*	-

p < .05

guiding this study. Statistical procedures and outcomes are reported and briefly discussed. An in-depth discussion of the results as they pertain to the research question can be found in Chapter 5.

Calibration of Comprehension. Both gamma and Pearson product moment correlation coefficients were calculated for each subject, correlating the confidence with the proportion correct across the 12 test items. Results of the 3 x 2 x 2 MANOVA with two dependent variables show a significant main effect for domain knowledge [$F(2, 50) = 10.93, p < .05$, effect size = .304, power = .99]. Univariate tests were also significant for both gamma [$F(1, 51) = 22.06, p < .05$, effect size = .301, power = .99] and Pearson product moment correlations [$F(1, 51) = 20.67, p < .05$, effect size = .288, power = .99]. Combined means and standard deviations for domain knowledge are reported in Table 18. No other significant main effects or interaction effects were detected. However, the means of the unguided questioning condition, although not significantly different, were higher than both the generic question stem and signal word conditions on both the gamma and Pearson product moment correlations (i.e., gamma = .26, .12, and .09, respectively; Pearson = .25, .16, and .10, respectively).

Confidence ratings for each subject were also analyzed. Results of the Factorial ANOVA show a significant main effect for domain knowledge [$F(1, 51) = 16.35, p < .05$, effect size = .243, power = .98]. As shown in Table 19, high domain participants mean of confidence rating was significantly higher than the mean of their other low domain knowledge counterparts. No other significant main effects or interaction effects were detected. However, the confidence rating of the generic question stem condition, although not

Table 18

Means and Standard Deviations for Domain Knowledge: Gamma and Pearson Product Moment Correlation Coefficients

Domain Knowledge	n	Gamma	Pearson Product Moment
High	34	-.03 (.39)	.01 (.35)
Low	29	.38 (.30)	.37 (.30)

p<.05

Note. Values enclosed in parentheses represent standard deviations.

Table 19

Means and Standard Deviations for Domain Knowledge: Confidence Rating and Post-Test Total Score

Domain Knowledge	n	Confidence Rating
High	34	57.6 (8.4)
Low	29	47.7 (10.3)

p<.05

Note. Values enclosed in parentheses represent standard deviations.

significantly different than either the signal word and unguided questioning conditions, was the highest.

Therefore, significant domain knowledge results in the post-test (see Table 7) suggest that while high knowledge participants scored significantly higher than their low knowledge counterparts, they were also more confident of their future performance on those questions. Nevertheless, when taking gamma and Pearson product moment correlations into consideration, low knowledge participants were more likely to accurately judge the level of their future performance.

Summary of Results

The results of this study are summarized in Table 20. They suggest that domain knowledge is an important factor when reading text. In most cases, high domain knowledge participants significantly outperformed their low domain knowledge counterparts. Other significant results included two-way and three-way interaction effects. In the case of the Post-Test, there was evidence that when reading low coherence text, the use of generic question stems was significantly more effective than using no questioning strategy at all on overall test performance as well as on textbase and elaborative inference questions.

In addition to domain knowledge main effects, there were main effects for treatment and text coherence on the summary recall variables. Significantly more unguided questioning participants cited propositions than any other treatment group while significantly more generic question stem participants cited elaborations. In regards to text coherence, significantly more participants reading low coherence text reordered the text while those reading high

coherence text cited significantly more macrolevel propositions. There were also three-way interactions concerning reorderings and Level 1 or topic/conclusion analysis suggesting that low coherence text, with various combinations of domain knowledge, helped the unguided and signal word participants to significantly outperform their high coherence counterparts with similar domain knowledge combinations. Furthermore, signal words helped the low coherence/low domain knowledge participants cite significantly more often the text topic compared to their unguided questioning/signal word counterparts with various combinations of domain knowledge.

In regards to the sorting data, there was significant improvement from pre- to post-treatment. Interestingly, however, the less structured the treatment, the greater the improvement. Thus, the unguided questioning participants performed significantly better than their signal word and generic question stem counterparts, in that order. The more structured the questioning strategy, the more interference when sorting the category words to resemble the scoring matrix.

Finally, while high domain knowledge participants were more confident in their post-test response accuracy, they were less likely to be accurate. Low domain knowledge participants had significantly higher calibration of comprehension scores despite their lack of confidence and poorer test scores.

In summary, there are a number of interesting findings from this study suggesting that text coherence, domain knowledge, and questioning strategies individually and collectively impact text comprehension.

Table 20

Summary of Results

Measure	Dependent Variable	Results
Post-Test	Post-Test Total Score	<ul style="list-style-type: none"> - significant main effect for domain knowledge favouring the high domain knowledge participants - significant two-way interaction between treatment and text coherence - low coherence text/generic question stem participants significantly outperformed the low coherence text/unguided questioning participants while the high coherence text/unguided questioning participants outperformed their low coherence text/unguided questioning counterparts.
	Textbase, Bridging Inference, Elaborative Inference, Problem Solving	<ul style="list-style-type: none"> - significant main effect for domain knowledge favouring the high domain knowledge participants on all dependent variables - significant two-way interaction between treatment and text coherence - significant univariate interactions for textbase and elaborative inference questions - for textbase questions, the low coherence text/generic question stem participants significantly outperformed the low coherence text/unguided questioning participants while the high coherence text/unguided questioning participants outperformed the low coherence text/unguided questioning participants and the high coherence/signal word participants - for elaborative inference questions, the low coherence text/generic question stem participants significantly outperformed the low coherence text/unguided questioning participants and the high coherence text/generic question stem participants

Table 20
Continued

Measure	Dependent Variable	Results
Calibration of Comprehension	gamma/Pearson product moment correlation	- significant main effect for domain knowledge favouring the low domain knowledge participants
	confidence ratings	- significant main effect for domain knowledge favouring the high domain knowledge participants
Sorting Data	pre- and post-treatment sorting scores	<ul style="list-style-type: none"> - significant main effects for sorting data - significant improvement from pre- to post-treatment - significant interaction effects for sorting data by treatment - significant multivariate interactions indicated that the unguided questioning participants significantly outperformed both their signal word and generic question stem counterparts.
Summary Recall a. Inference analysis	propositions	<ul style="list-style-type: none"> - significant main effects for domain knowledge favouring high domain knowledge participants - significant main effects for treatment indicating that the unguided questioning participants performed significantly better than the signal word participants
	elaborations, generalizations, and reorderings	<ul style="list-style-type: none"> - significant main effects for treatment with univariate main effects for elaborations indicating that the generic question stem participants significantly outperformed the unguided questioning participants and the signal word participants - significant main effects for text coherence with univariate main effects for reorderings favouring the low coherence participants - significant three-way interaction between treatment, text coherence, and domain knowledge with univariate effects for reorderings indicating that high domain knowledge/low text coherence/unguided questioning participants, high domain

Table 20
Continued

Measure	Dependent Variable	Results
Summary Recall b. Levels analysis	Levels 1, 2, 3, and 4	knowledge/high text coherence/signal word participants, and low domain knowledge/low text coherence/signal word participants had significantly more reorderings than their counterparts in those same conditions with various combinations of domain knowledge
		<ul style="list-style-type: none"> - significant main effects for domain knowledge with univariate main effects with Levels 3 and 4 favouring the high domain knowledge participants - significant main effects for text coherence with univariate main effects for Levels 2 and 3 favouring high coherence text participants - significant three-way interaction between treatment, domain knowledge, and text coherence with a significant univariate effect for Level 1 indicating that signal words assisted the low domain knowledge/low text coherence participants to state significantly more often what the text was about when compared to their signal word/unguided questioning counterparts with various combinations of domain knowledge/text coherency .

CHAPTER FIVE

DISCUSSION

Limitations of the Study

Prior to discussing the results of this study, some limitations must be clarified. When a study accesses a single class for participant screening and selection purposes, issues concerning data collection procedures, screening/selection criteria, and sample sizes ultimately manifest. Moreover, due to departmental policies for experimentation and experimental precedents, there are several time constraint issues which arise.

With respect to the screening and selection of participants, only one first year psychology class was approached. That same day, the class was screened for three other experiments. Students were given a testing package and were asked to complete all the enclosed instruments by the end of the class period and in the order that they were packaged. Students sat side-by-side in desks and occasionally asked their fellow classmates for clarification, despite being reminded to address the researchers who were supervising the testing period. Thus, student disregard for procedure as well as the potential for cheating and/or minimal effort was always present.

The primary purpose for screening was to ascertain levels of domain knowledge. In doing so, students were asked personal questions such as their relationships with those with heart disease, academic and non-academic courses taken concerning heart disease, and so forth. The expectation was that

students would be honest and forthcoming. Questions were also asked to determine in a quick and expedient manner specific domain knowledge. As with any questionnaire, it is possible that the respondent knows more than what is being asked of him or her. This is especially the case when the instrument is designed to be brief. Therefore, it is possible that each participant's score was not a true measure of his or her knowledge of heart disease. This may be compounded through the median-split classification procedure.

According to Voss and Bisanz (1985), low domain knowledge scores should range between 10 and 25% while high domain knowledge scores should range between 75 and 90%. Because of the small sample size from which to select participants, such grouping distinctions were not possible. Unfortunately, no other information regarding domain knowledge classification could be found in the literature. As a consequence, the median-split procedure was adopted. However, without confidence intervals, it is difficult to state definitively that a score which fell immediately above or beyond the median is indicative of high or low domain knowledge, respectively.

Caution must also be exercised when interpreting and generalizing the results of this study given its small, uneven cell sizes. Scores were analyzed by treatment, domain knowledge, and text coherence. This resulted in small numbers of participants in each cell; particularly in the case of interaction analyses. Such small, uneven numbers were beyond the researcher's control. As previously discussed, 193 first year psychology students were in attendance the day of the screening procedure. Of those students who were screened, 80 consented to partake in the second phase of the study. Once contacted, only 63 students participated. The remainder either did not consent to, or attend the second phase of the study due to lack of interest, completion of experimentation

credits, or absence.

Not only were there time constraints for the screening questionnaire (i.e., no longer than 10 minutes), there were also constraints imposed on the second portion of the study. The experiment was no longer than 120 minutes. All instructions and tasks were scripted as well as timed using a stopwatch. Students were to receive one research credit per half an hour. The Department of Psychology permitted the researcher to give two credits only. Any remaining experimental time was to be paid for by the researcher at a cost of five dollars per hour. Therefore, all participants of the second phase of the experiment received two credits and five dollars. In addition, it appeared that a two hour experiment was unprecedented. Getting students to commit to participation was problematic. Because of these departmental and participatory impediments, the amount of time accorded each task was affected. In some cases students may have been rushed to complete their tasks (e.g., summary recall and post-test) as well as prevented from learning and practicing the strategies involved in their treatment conditions. Had more time been available, a larger number of results might have confirmed the expectations.

The remainder of this chapter will provide conclusions for the research questions which guided this study, a synthesis of the findings, recommendations for educational practice, and suggestions for future research.

Conclusions Relative to the Research Questions

(1) In spite of text coherency and domain knowledge combinations, can we expect generic question stems to lead to greater measurable outcomes on memory (i.e., textbase model) and learning (i.e., situation model) variables as compared to signal words and/or unguided questioning techniques? In

accordance with previous findings (e.g., King, 1989), it was anticipated that those participants using generic question stems would outperform their counterparts using signal words and unguided questioning strategies. These results manifested either as a trend or as a significant finding.

As previously indicated, several variables were utilized in this study. The results generated by the variables suggest that generic question stems may be more effective for text memory and learning than any other questioning strategy. Based upon an earlier discussion in Chapter 2, such outcomes were expected. Rosenshine, Meister, and Chapman's (1996) review of questioning and comprehension studies concluded that generic question stems are the most successful prompt for facilitating the reading of expository text by obtaining the highest overall median effect size on experimenter-developed tests (87th percentile) as compared to signal words, story grammar, main idea, and question types. However, while the generic question stem participants were expected to outperform their counterparts in other treatment conditions, they were expected to do so with significant results.

Studies conducted by King (e.g., 1989, 1990c, 1992b, 1994b) were utilized in Rosenshine et al.'s review. Similar to this study, King assessed the impact of generic question stems on knowledge acquisition and retention. As with all of her previously cited studies, reported results were significant. Why the results of this study failed to be significant may be threefold. Firstly, King utilized a cooperative learning approach. Participants worked in small groups and posed and discussed questions with one another. For this study participants were required to work on an individual basis. This condition may have diminished opportunities for cognitive conflict and constructivist learning, and in turn, significant findings.

Secondly, King's work was largely restricted to lecture/lesson comprehension. This suggests that King's participants were adept at verbal language skills and listening comprehension. In comparison, this study required university students to read expository text. In doing so, we made the assumption that participants are entering the experiment with strong reading skills (i.e., strong vocabulary skills and reading speed and accuracy). If, however, such skills were lacking, experimental outcomes may have been affected. Because reading competency was not determined, it is difficult to know whether the quantity of significant outcomes would have remained the same.

Thirdly, as with all of questioning strategies reviewed by Rosenshine et al., King's treatment ranged from 4 to 8 sessions. This study was restricted to 120 minutes for instruction, treatment, and testing. This may have had an adverse impact on strategic acquisition. Therefore, given these considerations, it is difficult to compare the outcomes generated by King's research to those produced by this study. Had King's conditions been replicated for this study, more significant results using generic question stems may have manifested.

Although King did not utilize the same variables and measures as those employed for this study (i.e., sorting data, summary recall, and calibration of comprehension), her variables and measures were similar as they assessed knowledge acquisition, knowledge mapping, and comprehension. Due to these similarities, it was assumed that generic question stems would produce similar outcomes. While they predominately did, few results were significant. The following examination of each instrument will provide greater insight.

The Post-Test was designed to assess knowledge acquisition and retention. This objective was comprised of an assessment of global knowledge

derived from the text as well as the delineation between textbase and situation model development. Separate analyses were conducted using general test scores and scores for question types (i.e., bridging inference, elaborative inference, textbase, and problem solving). While there were no treatment main effects from either analyses, the generic question stem condition did generate higher means on all variables.

The Sorting Data measure is identical to that used by McNamara et al. (1996). Its purpose is to assess changes in a reader's conceptual structure and the amount learned as a result of reading a text. Consequently, it is suitable for assessing situation model understanding. It was included in this study to demonstrate how questioning strategies, particularly generic question stems, bolster conceptual understandings and generate links within text and to previous knowledge. The results were rather intriguing. Prior to treatment, there were no significant differences between conditions on the pre-treatment sorting results. However, following treatment, there were significant changes favouring the unguided questioning condition. Its participants significantly outperformed both the signal word and the generic question stem conditions. Upon closer examination of the post-treatment sorting means, the less structured the questioning technique, the better the sorting score. One possible explanation may be that structured questioning techniques cause interference in conceptual understanding. In being asked to follow specific questioning formats, participants may have been obstructed from making expert-like connections between concepts. That is, the more structured techniques may have diverted one's attention from the conceptual underpinnings of the text or taxed vital cognitive resources necessary for conceptual understanding.

A second possibility is that the more structured the questioning strategy,

the greater the number of perceived connections. Examining the combined means of both sorts, the generic question stem participants showed no improvement from the pre-treatment sort to the post-treatment sort. Why was there no improvement? To answer this question it is necessary to examine the scoring procedure for this measure. According to McNamara et al. (1996), the score is a measure of harmony between the participant's sorting matrix and a weight matrix of an ideal sorting. The weight matrix is comprised of 8 categories of related items from the text. Five categories are assigned a weight of 1 and 3 categories are assigned a weight of .5. All other sorts are assigned a negative value (-.0456204). If a word is placed into a category in which it does not belong, it will receive a negative value. The more incorrect categorical connections, the lower one's sorting score. In the case of the generic question stems, this strategy requires the reader to make connections both within the text and to his or her previous knowledge base. In doing so, it is possible that a liberal number of connections amongst concepts become visible in spite of the more conservative ideal matrix. Thus, the participants in the generic question stem condition may have derived connections between concepts which had intuitive and medical merit. They perceived connections which do not exist according to the scoring matrix, and were consequently incorrect. However, this does not mean that they were all incorrect, only that they did not meet scoring criterion.

All participants were instructed to justify their categories. From an observational perspective, those in the generic question stem condition created a greater number of categories and cited relationships which often exceeded the parameters of the heart disease text. Thus, being instructed to generate questions which necessitate connections across the text as well as to previous

knowledge (i.e., previous course work or personal experience, connections with the text entitled "Blood", and so forth) may have lead the generic question stem participants to make connections which weakened their sorting score rather than improving upon it. The strongest sorting score belonged to the unguided questioning condition. Because they were not asked to make connections when generating questions, they may have been better able to see the most obvious relationships amongst the concepts. Therefore, to suggest that generic question stems and signal words fail to help readers see conceptual relationships may be premature, particularly when taking their performance on other variables and measures into consideration. That is, if structured questioning strategies were as detrimental as the sorting data suggests, would not the performances of the generic question stem and signal word participants be equally poor on the Post-Test and Summary Recall variables and measures? Since this is not the case, it appears that the more connections a questioning strategy requires a reader to make within text and to previous knowledge, the more connections will be made to the detriment of one's sorting score.

The Summary Recall measure was included in this study to assess situation and textbase model development. Assessment was accomplished through macrostructure levels analysis (i.e., macropropositions and text details) and text reproduction and inferencing (i.e., propositions, elaborations, generalizations, and text reorderings). Again, it was expected that those within the generic question stem condition would have superior scores on all independent variables, particularly those pertaining to situation model development. It should be noted that generalizations are exemplary indicators of a situation model with elaborations, reorderings, and macropropositions (i.e.,

Levels 1, 2, and 3) typifying a less well-developed situation model with underlying textbase characteristics. Level 4 details are exclusive indicators of a textbase model representation.

With respect to the recall of propositions, the participants in the unguided questioning condition recalled significantly more propositions than those in the signal word condition, but not significantly more than their generic question stem counterparts. Although unexpected, the superiority of the unguided questioning condition is logical. Like the Sorting Data outcomes, unguided questioning participants were not required to engage in extraneous question generation. Thus, they were better able to concentrate upon the details of the experimental text as opposed to making connections between concepts and to prior knowledge.

The generic question stem participants did receive the highest combined means for generalizations and reorderings and performed significantly higher than either the signal word or unguided questioning condition on elaborations. These results suggest that generic question stems assist the reader in processing primary ideas from the text, linking them with previous knowledge, and organizing them in a manner which is personally meaningful. Thus, generic question stems expedite textbase and, to a greater extent, situation model development.

Although the performance differences were marginal, the generic question stem participants cited Level 1 (i.e., Topic/Conclusion) and Level 3 (i.e., Subheadings) responses more frequently than any other condition. While this suggests that structured questioning incites macrolevel processing, the differences on all four Levels are too inconsistent to denote a pattern between conditions.

Therefore, it appears that generic question stems typically lead to greater measurable outcomes than signal words and/or unguided questioning techniques. As discussed above, results which merely indicated a trend may have been significant had participants been allowed to discuss their questions and answers with fellow condition members and/or been given more time to engage in question training and practice. Nevertheless, there is evidence to suggest that generic question stems had a positive impact on knowledge acquisition, retention, and transfer. More simply, they assisted the university-age participants in developing textbase and situation model representations.

(2) Will there be interaction effects on these variables suggesting that combinations of domain knowledge and text coherency are uniquely affected by generic question stems, signal words, or unguided questioning? Previous research conducted by Kintsch and colleagues (e.g., McNamara & Kintsch, 1996; McNamara et al., 1996) has involved the impact of domain knowledge and text coherency on expository text processing. This study has included both of these factors in addition to questioning as a treatment condition.

Like Kintsch and colleagues, domain knowledge was divided between high and low knowledge while text coherency was manipulated to include high text coherency (i.e., good macrostructure/good microstructure) and low text coherency (i.e., poor macrostructure/poor microstructure). According to the findings of Kintsch and colleagues, high domain knowledge participants perform optimally when given low coherency text while low domain knowledge participants perform best with high coherency text. Given these outcomes, the purpose of this study was to better understand these results and to determine whether the inclusion of specific questioning strategies replicated or altered

these findings. Simply, do generic question stems, signal words, and/or unguided questioning uniquely interact with domain knowledge and text coherency, alone or in combination, when reading expository text?

As concluded through this study, specific questioning strategies do interact with domain knowledge and text coherency, both alone and in combination. However, it should be noted that the most common significant outcomes involved domain knowledge main effects. High domain knowledge participants significantly outperformed their low domain knowledge counterparts on the Post-Test total score, all four Post-Test question types (i.e., textbase, bridging inference, elaborative inference, and problem solving), Summary Recall propositions, and Summary Recall Levels 3 and 4. Significant main effects for text coherence were also detected on the Summary Recall reorderings favouring the low text coherence participants and on the Summary Recall Levels 2 and 3 favouring the high text coherence participants.

As to why high domain knowledge participants performed so well on the above-cited dependent variables, it may be that prior knowledge assists the reader in developing a strong textbase model, and to a lesser extent, a situation model. As suggested by Kintsch and colleagues (McNamara & Kintsch, 1996; McNamara et al., 1996), with the exception of low domain knowledge readers given low coherence text, all readers can form an adequate textbase and recall text quite well. After all, a textbase model is indicative of text memory. Thus, a reader should be able to successfully recall propositions without necessarily understanding the relations between them. It is not until they are forced to engage in active processing (i.e., high knowledge readers with low coherence text) that situation model development is expected to take place (McNamara & Kintsch, 1996; McNamara et al., 1996).

However, with the exception of Summary Recall reorderings, which will be discussed at a later point, the findings of Kintsch and colleagues were not replicated in this study. High domain knowledge participants who read low coherence text did not demonstrate significantly superior situation model development or superior learning. Rather, a main effect of high domain knowledge led to significantly superior textbase model development when compared to any other combination of specific questioning strategies and/or text coherency. On the basis of this outcome, we may conclude that those readers with high domain knowledge are most advantaged in regards to significant textbase model development; particularly when time for strategic instruction and acquisition is limited. They are able to encode text information into already existing schemas by way of selective combination (Sternberg, 1985) or internal connections (Mayer, 1989, 1992), but appear less inclined to engage in selective comparison (Sternberg, 1985) or external connection generation (Mayer, 1989, 1992).

So why did the combination of high domain knowledge with low text coherency fail to generate significant situation model outcomes? Perhaps the answer involves the inclusion of questioning strategies. That is, too much cognitive capacity and processing time may have been subsumed by question generation, which in turn interrupted the active processing of text. By being instructed to generate questions in such a limited period of time, the attention required for filling-in the gaps of the low coherence text may have been diverted, not enhanced. Therefore, instead of eliciting the connections necessary for text clarity, the combination of high domain knowledge and question generation incited participants to process information which matched their existing knowledge representations. In other words, by way of question

generation, high domain knowledge participants were reinforcing their textbase representations at the expense of a situational representation.

Results concerning text coherence appear to be straightforward. Participants reading low coherence text engaged in more Summary Recall reorderings than those reading high text coherence text. Without the structure imposed by high coherence text, participants tended to reorder the text during Summary Recall. That is, they were less likely to recall text paragraphs and details in the order in which they were presented compared to those participants reading high coherence text. Reasons for this outcome may be twofold. Firstly, reorderings primarily fall under the domain of situation models. That being the case, it may be that readers engaged in more active processing when reading the low coherence text. They were compelled to see connections between paragraphs and details causing them to reorder the information during recall. However, if this were the case, we would expect to see significant results on other situation model variables such as bridging inference questions, sorting data, generalizations, and so forth.

A more plausible theory may be that those participants reading low coherence text had a difficult time recalling the text. This in turn lead to the inadvertent reordering of paragraphs and details. Instead of information being reordered for the purposes of clarity and personally meaningful connections, it was recalled randomly.

In regards to the Levels outcomes, those participants reading high coherence text recalled significantly more Level 2 and Level 3 data. Therefore, they recalled subtopics and subheadings significantly more often than their low text coherence counterparts. This can be easily explained since the high coherence text explicitly marked the subtopics comprised in Level 2.

Furthermore, the Level 3 subheadings were easier to interpret in the high coherence text due to their explicitness at the local, and even global levels. Therefore, while there were significant main effects involving domain knowledge and text coherency, they did not overshadow the interaction effects.

In regards to the interaction effects, several interesting results were detected. For instance, there was a significant two-way interaction between treatment and text coherence for the Post-Test total score as well as textbase and elaborative inference questions. In the case of the Post-Test total score data, it appears that when reading low coherence text, generic question stems significantly improved overall test performance as compared to using no questioning strategy at all. Generic question stems combined with low coherence text produced the highest mean of all combinations of text coherence and treatment. This was followed by unguided questioning with high coherence text, which was also significantly higher than unguided questioning with low coherence text. Therefore, it may be that generic question stems induce more active text processing than any other combination of text coherence and treatment. Why these results were not replicated for the generic question stem/high coherence text participants is unclear. Perhaps when text is better structured and easier to read, the absence of a structured questioning strategy is more productive as there is less cognitive capacity being consumed. Consequently, more attention is being diverted to text comprehension versus strategy manipulation.

Similar results manifested on textbase and elaborative inference questions. In regards to the textbase questions, the generic question stem/low coherence text participants significantly outperformed their unguided questioning/low coherence text counterparts. They also had the highest overall

mean on the textbase questions with the combination of text coherence and treatment. Furthermore, the unguided questioning/high coherence text participants significantly outperformed their low coherence counterparts as well as the signal word/high coherence participants. For elaborative inference questions, the generic question stem/low coherence participants significantly outperformed the unguided questioning/low coherence text participants in addition to their generic question stem/high coherence text counterparts. Finally, the generic question stem/low coherence participants produced the highest overall mean on elaborative inferencing questions with the combination of text coherence and treatment. Therefore, as with overall test performance, generic question stems improve text comprehension when combined with low coherence text. Only in the case of textbase questions did unguided questioning/high coherence text prove to be a significant asset. On the basis of these outcomes we may conclude that generic question stems combined with low coherence text improve Post-Test performance, particularly those questions reflecting a textbase model representation.

A significant three-way interaction was also detected between treatment, text coherence, and domain knowledge on the Summary Recall reorderings as well as on Level 1 (i.e., topic and conclusion) citations. In regards to reorderings, those participants with high domain knowledge/low text coherency/unguided questioning generated significantly more reorderings than any other combination of text coherency, domain knowledge, and unguided questioning as well as the low domain knowledge/high text coherency and high domain knowledge/low text coherency members of the signal word condition. This was followed by the significantly superior performances of the high domain knowledge/high text coherency/signal word participants and the low domain

knowledge/low text coherency/signal word participants over the majority of their signal word and unguided questioning counterparts.

As previously discussed, Summary Recall reorderings are indicative of a less than well-developed situation model. Therefore, the superior performance of the high domain knowledge/low text coherency participants is the only outcome which approximates the findings of Kintsch and colleagues in regards to significant situation model outcomes (McNamara & Kintsch, 1996; McNamara et al., 1996). As discussed above, the inclusion of a questioning strategy with high domain knowledge and low text coherency may interrupt the development of a strong situation model representation. Because this is the only significant outcome involving this combination with questioning, specifically unguided questioning, it may be true that the inclusion of question generation interferes with the situation model development of high domain knowledge/low text coherency participants. This does not appear to be the case with other combinations of domain knowledge and text coherency. For instance, signal word participants with high domain knowledge/high text coherency and low domain knowledge/low text coherency did significantly better than the majority of their counterparts including those with high domain knowledge/high text coherency in the unguided questioning condition. Thus, questioning, more specifically structured questioning appears to assist those reading under either optimal conditions or meager conditions. In keeping with previously cited outcomes, this is especially true for those participants with the fewest advantages (i.e., low domain knowledge and/or low text coherency).

Similar outcomes arose with the Level 1 results. Those signal word participants with low domain knowledge/low text coherency, low domain knowledge/high text coherency, and high domain knowledge/high text

coherency significantly outperformed the majority of the signal word/unguided questioning participants. The only exception was the performance of the low domain knowledge/high text coherency participants in the unguided questioning condition whose scores tied with their low domain knowledge/high text coherency/signal word counterparts. Therefore, combined with previously cited results, it may be that structured questioning strategies and high domain knowledge are more reliable indicators of adaptive textbase and situation model development than the combination of high domain knowledge and low text coherency alone.

(3) Will the employment of generic question stems lead to superior calibration of comprehension scores than those generated by signal words and/or unguided questioning conditions?

One of the most important elements of learning from text is the reader's ability to assess his or her level of comprehension.

Unfortunately, it has been shown that the ability to judge whether one has comprehended read material is lacking. While the cause is presently unknown, previous research has suggested that domain knowledge plays a significant role (e.g., Glenberg & Epstein, 1987; Glenberg, Sanocki et al., 1987).

Specifically, it appears that high domain knowledge participants miscalculate their future test performance. As a result, their subjective assessments of knowledge are inaccurately inflated in comparison to their performance on an objective measure.

Other factors contributing to poor calibration include text structure, text genre, and age-appropriateness. According to research conducted by Weaver and Bryant (1994), and Weaver, Bryant, and Burns (1995), text which is readable and age-appropriate leads to higher calibration scores. Moreover,

depending upon the text genre (i.e., expository or narrative), different measures generate different calibration results. Typically, narrative text incites thematic-based processing while expository text induces detail-oriented processing. Readers of expository text tend to monitor their performance on detail-oriented measures with greater accuracy than on thematic-oriented measures. These outcomes are reversed for readers of narrative text. However, there is evidence to suggest that when additional processing is required, calibration scores improve (e.g., Glenberg, Sanocki et al., 1987; Maki et al., 1990; Maki & Swett, 1987). This study sought to investigate these results and examine the impact of additional processing on calibration of comprehension. Moreover, it explored whether the inclusion of questioning strategies increased reader self-efficacy by way of confidence ratings.

Adhering to the theories and outcomes of previously cited research, this study paid careful attention to age-appropriateness and text genre considerations. For instance, depending upon the Readability Index observed (see Table 1), the experimental text entitled "Heart Disease" falls between the categories of "easy" (below grade 8) and "standard" (around grade 12) (Weaver & Bryant, 1994). The results of Weaver and Bryant (1994) suggest that for university age participants, "standard" text should be used to ensure higher calibration scores. However, the results of this study did not replicate this outcome. Furthermore, the experimental text was manipulated to represent both high coherency and low coherency. According to previous research, the more readable the text, the better the calibration score. Again, this outcome did not manifest. In fact, there were no text coherency main effects suggesting that neither high nor low coherency text manipulations affected calibration.

Thirdly, because the experimental text can be characterized as

expository, the Post-Test was constructed to ascertain detail. Although there were no theme-based results with which to compare, the absence of high calibration amongst all participants suggests that text genre and the manner in which it is tested (i.e., detailed-oriented test design) do not conclusively improve calibration scores. And finally, the inclusion of structured questioning did not provide the additional processing necessary to improve calibration of comprehension. Participants using generic question stems did not out perform their signal words or unguided questioning counterparts.

The only significant outcome of this analysis was a main effect for domain knowledge. In keeping with previous research, those participants with high domain knowledge were over-confident when judging their future performance on specific Post-Test questions. Thus, low knowledge participants were more accurate in their judgments. According to Glenberg, Sanocki et al., (1987) and Glenberg and Epstein (1987), high domain knowledge participants typically base their judgments on topic familiarity as opposed to text comprehension. Despite the inclusion of a questioning strategy, high domain knowledge participants maintained an inaccurate perception of subject knowledge. The task of generating questions failed to alert them to potential knowledge gaps. Therefore, it appears that knowledge alone influences calibration of comprehension. Neither text coherency nor treatment had any known impact.

Synthesis

Throughout this chapter, several important outcomes have been discussed, each in relation to the questions which guided this study. This section will attempt to synthesize those findings, bearing in mind the goals of the experiment.

This study was founded on the need to better understand the conditions necessary for adaptive expository text processing through questioning. In doing so, attempts were made to simulate the elements typified in the majority of reading situations. That is, levels of domain knowledge were ascertained for each participant while text coherency was manipulated to reflect the most and least advantageous reading contexts.

These reading environment combinations were previously encountered in several studies conducted by Kintsch and colleagues (McNamara & Kintsch, 1996; McNamara et al., 1996). In an attempt to improve instructional text quality, the researchers focused on the effects of reader's background knowledge and text coherence. Specifically, they sought to determine whether text memory and learning are enhanced or diminished by improving text comprehensibility. As previously discussed, the answer to this question was complex. Depending upon one's level of domain knowledge, text either had to challenge or alleviate the reader's cognitive resources. But is such tailoring feasible? Can expository text be appropriately matched with a reader's individual needs? Might there be a strategy which circumvents individualized text assignment ? One approach may be through questioning. Perhaps strategic questioning can invoke the additional processing required to make text both challenging and comprehensible, in spite of one's level of domain knowledge.

A number of studies have shown that comprehension can be significantly improved through strategic questioning and expository text (e.g., Andre & Anderson, 1978-1979; Davey & McBride, 1986; Lysynchuk et al., 1990). The most successful strategy involved generic question stems through Reciprocal Peer Questioning (e.g., King, 1989). In using the stems, participants were

required to draw inferences about new information, to take a new perspective on existing knowledge, and to elaborate upon new material and existing knowledge structures. Simply, the participants were required to generate the internal and external connections necessary for text memory and learning. Unfortunately, generic question stems were tested exclusively with lesson/lecture comprehension and in a cooperative learning environment. This study has been an opportunity to observe the reliability of question stems when used on an individual basis as well as in conjunction with expository text.

The Kintsch model of text comprehension (e.g., Kintsch, 1988; Kintsch & van Dijk, 1978) was employed for this study. The construction-integration (CI) model is connectionist model. Through the construction and integration of a network of concepts, a mental representation of the text is achieved. The most important components of this representation include textbase and situation models. Therefore, to determine whether comprehension occurred in this study, it was necessary to assess both models. Like Kintsch and colleagues (e.g., McNamara & Kintsch, 1996; McNamara et al., 1996), this involved a sorting task, a post-test, and a recall measure. Because text comprehension also involves self-monitoring, a calibration of comprehension measure was included. The results demonstrated the salience of domain knowledge and questioning, but they did not replicate the findings of previous research. That is, there were fewer interaction effects than discovered by Kintsch and colleagues and fewer main effects for questioning as ascertained by King (e.g. King, 1989).

As suggested, this study did not replicate the experimental contexts of either King or Kintsch and colleagues. As with previous studies, questioning strategies were not taught and practiced over several sessions, nor were they employed in a cooperative setting (e.g., King, 1989). Furthermore, text

comprehension was assessed without the involvement of a specific reading strategy. Therefore, it was unlikely that the same results generated by previous researchers would manifest. Nevertheless, the outcomes of this study proved to be promising. Although not always significant, generic question stems appear to provide the additional processing required for text memory and text learning; particularly for those readers lacking domain knowledge and/or reading low coherence text.

As indicated in an earlier discussion, the most common significant results involved domain knowledge main effects. High domain knowledge participants significantly outperformed their low domain knowledge counterparts on primarily textbase model variables in addition to calibration of comprehension and confidence ratings scores. McNamara and Kintsch (1996) similarly detected domain knowledge main effects favouring high domain knowledge participants on propositional recall, sorting scores, and multiple-choice question responses (i.e., textbase and bridging inference) in Experiment 1. When domain knowledge was changed from a dichotomous to continuous variable in Experiment 2, there was a marginal main effect only on open-ended question responses (i.e., textbase and bridging inference). A similar outcome was also detected by McNamara et al., (1996) with high domain knowledge participants significantly outperforming their low domain knowledge counterparts on post-test questions only (i.e., textbase, bridging inference, elaborative inference, and problem solving). Unfortunately, no information indicating whether the significant results involved questions which were textbase or situational was provided in either study. However, given how few situation model questions were included in the post-tests, and that no other situation model measures generated significant domain knowledge main effects, it appears that high

domain knowledge participants excelled on textbase model variables. If this assumption is correct, the findings of Kintsch and colleagues support the theory proposed by this study. That is, in contrast to text learning, text memory and text recall generate the greatest number of significant results when time is limited for strategy involvement and acquisition. Compared to the cognitively taxing process of schematic accommodation, those with the greatest knowledge bases can simply encode or assimilate text information into already existing schemas.

In regards to situation model variables, few significant outcomes arose in this study. Was this a consequence of time constraints, the nature of the questioning strategies, or a combination of both? When considering the outcomes of Kintsch and colleagues, we may conclude that the inclusion of questioning strategies with such limited instructional time is detrimental to situation model development. Given approximately the same reading time as with this study, McNamara and Kintsch (1996) and McNamara et al. (1996) detected significant results on all situation model variables (i.e., sorting task, bridging inference questions, and problem solving questions). Furthermore, as predicted by the researchers, high knowledge readers performed better on these variables after reading the low coherence text while low knowledge readers did generally better after reading high coherence text. So why were these results not replicated for this study? Perhaps the inclusion of questioning strategies within a limited amount of time caused a processing interference. Instead of actively processing the text and acquiring a situational understanding, participants were preoccupied with question generation. This theory appears feasible in light of the three-way interaction on Summary Recall reorderings. Those participants with high domain knowledge reading low coherence text achieved optimal performance using an unstructured

questioning technique. Thus, the less cognitive capacity occupied by structured questioning, the better the participants performed on situation model variables. Additional support for this position can be found with the significant two-way interaction between treatment and text coherence on Post-Test total score and elaborative inference questions. Those participants reading high coherence text performed optimally when using the less imposing unguided questioning strategy. Therefore, in spite of a few significant effects involving less well-developed situational variables, the inclusion of structured questioning strategies during a short instructional period appears to have disrupted the active processing required for strong situational understanding.

Given that these findings contradict those of Kintsch and colleagues, can it be said that text tailoring is the only viable solution to adaptive text processing? Not when we consider the strong trend evidenced by generic question stems. While not always significantly superior, generic question stem participants outperformed all other treatment participants on the Post-Test total score; textbase, elaborative inference, bridging inference, and problem solving Post-Test questions; confidence ratings; Summary Recall generalizations, elaborations, and reorderings; and Level 1 and Level 3 citations. Moreover, generic question stems were involved with several significant interaction effects. This suggests that generic question stems are viable predictors of textbase and situation model development.

So why are generic question stems so beneficial? The answer may be found in the stems themselves. Each stem is designed to invoke textbase and situational understandings of the text. Such understanding is accomplished by explicitly guiding the reader to generate connections both within the text and to his or her previous domain knowledge. Simply, generic question stems engage

the reader in the active processing of text. What may have prevented more significant findings concerning generic question stems was the time restriction imposed on the experiment. Although not used for the purposes of text comprehension, King's (e.g., 1989) studies demonstrated that generic question stems result in significant outcomes. In all instances, generic question stems were practiced over several sessions. This suggests that in order for generic question stems to be maximally beneficial, sufficient time for acquisition and utilization must be given.

The final goal of this study was to determine whether the active processing induced by questioning can improve calibration of comprehension scores. No significant treatment effects were found, only significant main effects for domain knowledge. In keeping with previous research, high domain knowledge readers typically base their judgments of future performance upon topic familiarity as opposed to text comprehension. Perhaps the inclusion of questioning strategies did not provide the additional processing required for adaptive calibration or perhaps they failed to be effective due to time constraints. Had the participants been given more time to practice with and internalize the strategies, significant results may have arisen. Nevertheless, while the inclusion of questioning strategies failed to induce calibration of comprehension, generic question stems did foster reader confidence. On the basis of confidence means, there was evidence that generic question stem participants felt more confident of their future performance than their signal word and unguided questioning counterparts (i.e., 54.6, 50.9, and 53.4, respectively). Although these results are not significant, they do convey that generic question stems may increase reader self-efficacy. Again, future research is required to verify or refute this claim.

In summary, while it was previously believed that text must be tailored to the level of one's domain knowledge for active processing and situation model development, new findings suggest that this may not be necessary. Although not significant, generic question stems appear to elicit the internal and external connections for adaptive expository text processing. Moreover, they help incite the confidence necessary for reader self-efficacy amongst university-age participants.

Recommendations for educational practice and future research

Several recommendations for future educational practice and research evolved from the results and discussion.

- (1) To ensure the development of textbase, and especially situation model representations with university-age participants, it is necessary to provide additional processing in the way of structured questioning techniques. According to the findings of this study, the most advantageous questioning strategy concerns generic question stems. Although not always significant, results derived from generic question stem participants demonstrated a decisive trend toward superior textbase and situation model representations during expository text processing.
- (2) In lieu of strategic instruction, high domain knowledge is the most reliable indicator of textbase and situation model development amongst university-age participants. Thus, the provision of sufficient background information is necessary for adaptive text comprehension in the presence of time constraints and poor strategic processing.
- (3) Amongst university-age participants, text coherency has little impact upon text comprehension. Therefore, more attention should be diverted to the

development of domain knowledge and strategic instruction to ensure adaptive text processing.

(4) Although not significant, the utilization of structured questioning can influence reader self-efficacy. Amongst university-age participants, generic question stems provided the support necessary for reader confidence on future test performance.

(5) Sufficient practice must be given when introducing a new questioning technique. To ensure that internalization occurs, ample time must be provided for understanding and practice. Unfortunately, this experiment adhered to strict time constraints. Had more time been available, a larger number of results might have confirmed these expectations.

(6) Previous research on generic question stems has involved group learning. It may be that generic question stems are best learned in a cooperative setting. Therefore, future text processing research concerning the questioning strategies used in this study should include groups of 3 to 4 participants who are instructed to share and discuss their questions and answers. It may be that the cognitive conflict required for optimal text processing is more likely to occur in a cooperative ethos as compared to an individualized setting.

(7) Further research with adolescent populations is warranted. Research has typically focused upon primary and college-aged students with little attention directed toward adolescents. Such an absence can lead to the generation of conjectures regarding middle years development in lieu of evidence-based generalizations. Some conclusions derived from previous research suggest that adolescents who are trained to engage in question-generation strategies perform at a level of significance on comprehension and metacomprehension measures (Davey & McBride, 1986) as well as on products tests (Laidlaw, Skok,

& McLaughlin, 1993). Moreover, in comparison to their younger counterparts, adolescents are more likely to engage in directed and spontaneous transfer (Pressley & Dennis-Rounds, 1980) when generalized results and specific examples applicable to the transfer task are given (Gick & Holyoak, 1983), when the learning situation is analogous to the transfer situation (Gick & Holyoak, 1980), and when they are trained to individually monitor their own learning (Bender, 1986). Thus, the assumption that middle years students are developmentally equivalent to younger children and adults is unwarranted. Any future research concerning expository text processing and strategic questioning with adolescents is necessitated given their developmental differences and requirements.

(8) Future research evolving from this study should include the scoring and categorization of participant generated questions. The collection of participant generated questions during this study was to ensure that questions were being generated, that participants were following the instructions and requirements of their treatment condition, and to enforce the benefits of note-taking and written recording on memory and learning. At no time was it indicated that questions would be scored and categorized for analysis. Future research will include the categorization of questions according to textbase and situation model representations. This will provide additional insight into the benefits of specific questioning techniques on a reader's mental representation of text.

REFERENCES

- Adams, M. (1990). Beginning to Read: Thinking and learning about print. Cambridge, MA: The MIT Press.
- Albrecht, J.E., & O'Brien, E.J. (1993). Updating a mental model: Maintaining both local and global coherence. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 1061-1070.
- Albrecht, J.E., O'Brien, E.J., Mason, R.A., & Myers, J.L. (1995). The role of perspective in the accessibility of goals during reading. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 364-372.
- Alexander, P.A., Schallert, D.L., & Hare, V.C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. Review of Educational Research, 61, 315-343.
- Anderson, J.R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.
- Anderson, J.R., Reder, L. M., & Simon, H.A. (1996). Situated learning and education. Educational Researcher, 25, 5-11.
- Andre, M.E.D.A., & Anderson, T.H. (1978-1979). The development and evaluation of a self-questioning study technique. Reading Research Quarterly, 14, 605-623.
- Ausubel, D.P. (1968). Educational psychology: A cognitive view. New York, NY: Holt, Rinehart, and Winston.

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Bartlett, F. (1932). Remembering. Cambridge, UK: Cambridge University Press.

Baron, R.A., Earhard, B., & Ozier, M. (1998). Psychology (2nd Canadian ed.). Toronto, Ont: Allyn and Bacon.

Barton, S.B., & Sanford, A.J. (1993). A case-study of anomaly detection: Shallow semantic processing and cohesion establishment. Memory and Cognition, 21, 477-487.

Beck, I.L., & McKeown, M.G. (1989). Expository text for young readers: The issue of coherence. In L. Resnick (Ed.), Knowing, learning, and instruction: Essays in honor of Robert Glaser (pp. 47-66). Hillsdale, NJ: Erlbaum.

Beck, I.L., McKeown, M.G., & Gromoll, E.W. (1989). Learning from social studies texts. Cognition and Instruction, 6, 99-158.

Beck, I.L., McKeown, M.G., Sinatra, G.M., & Loxterman, J.A. (1991). Revising social studies text from a text-processing perspective: Evidence of improved comprehensibility. Reading Research Quarterly, 26, 251-276.

Bender, T. (1986). Monitoring and the transfer of individual problem solving. Contemporary Educational Psychology, 11, 161-169.

Bereiter, C. (1985). Toward a solution of the learning paradox. Review of Educational Research, 55, 201-226.

Bereiter, C. (1994). Constructivism, socioculturalism, and Popper's World 3. Educational Researcher, 23, 21-23.

Bereiter, C., & Scardamalia, M. (1987). The psychology of written composition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Black, J.B. (1985). An exposition on understanding expository text. In B.K. Britton & J.B. Black (Eds.), Understanding expository text (pp. 249-267). Hillsdale, NJ: Lawrence Erlbaum Associates.

Britton, B.K., Glynn, S.M., & Smith, J.W. (1985). Cognitive demands of processing expository text: A cognitive workbench model. In B.K. Britton & J.B. Black (Eds.), Understanding expository text (pp. 227-248). Hillsdale, NJ: Lawrence Erlbaum Associates.

Britton, B.K., & Gulgoz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. Journal of Educational Psychology, 83, 329-345.

Britton, B.K., van Dusen, L., Gulgoz, S., & Glynn, S.M. (1989). Instructional texts rewritten by five expert teams: Revisions and retention improvements. Journal of Educational Psychology, 81, 226-239.

Brown, A.L., & Campione, J.C. (1986). Psychological theory and the study of learning disabilities. American Psychologist, 41, 1059-1068.

Brown, A., & Palincsar, A. (1989). Guided, cooperative learning and individual knowledge acquisition. In L.B. Resnick (Ed.), Knowing, learning, and instruction: Essays in honor of Robert Glaser (pp. 393-451). Hillsdale, NJ: Lawrence Erlbaum Associates.

Carlson, N.R., & Buskit, W. (1997). Psychology: The science of behavior (5th ed.). Toronto, Ont: Allyn and Bacon.

Clifford, M.M. (1988). Failure tolerance and academic risk-taking in ten- to twelve-year-old students. British Journal of Educational Psychology, 58, 15-27.

Cobb, P. (1988). The tension between theories of learning and instruction in mathematics education. Educational Psychologist, 23, 87-103.

Collins, A., Brown, J.S., & Newman, S.E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L.B. Resnick (Ed.), Knowing, learning and instruction: Essays in honor of Robert Glaser (pp. 453-494). Hillsdale, NJ: Lawrence Erlbaum Associates.

Côté, N., Goldman, S.R., & Saul, E.U. (1998). Students making sense of informational text: Relations between processing and representation. Discourse Processes, 25, 1-53.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

Davey, B., & McBride, S. (1986). Effects of question-generation training on reading comprehension. Journal of Educational Psychology, 78, 256-262.

De Jong, T., & Ferguson-Hessler, M.G.M. (1996). Types and qualities of knowledge. Educational Psychologist, 31, 105-113.

Derry, S.J. (1990). Learning strategies for acquiring useful knowledge. In B.F. Jones & L. Idol (Eds.), Dimensions of thinking and cognitive instruction (pp. 347-379). Hillsdale, NJ: Lawrence Erlbaum Associates.

Derry, S.J. (1992). Beyond symbolic processing: Expanding horizons for educational psychology. Journal of Educational Psychology, 84, 413-418.

Dochy, F.J.R.C. (1996). Assessment of domain-specific and domain-transcending prior knowledge. In M. Birenbaum & F.J.R.C. Dochy (Eds.), Alternatives in assessment of achievements, learning processes and prior knowledge (pp. 227-265). Boston, MA: Kluwer Academic Publishers.

Dweck, C.S. (1986). Motivational processes affecting learning. American Psychologist, 41, 1041-1048.

Dweck, C.S., & Elliott, E.S. (1983). Achievement motivation. In P. Mussen (Ed.), Carmichael's manual of child psychology (Vol. 4, pp. 643-691). New York, NY: Wiley.

Engestrom, Y. (1987). Teaching dialectical thinking through model construction. In J. T.E. Richardson, M.W. Eysenck, & D.W. Piper (Eds.), Student learning: Research in education and cognitive psychology (pp. 120-132). Philadelphia, PA: Open University Press.

Estes, W.K. (1995). A general model of classification and memory applied to discourse processing. In C. Weaver, III, S. Mannes, & C.R. Fletcher, (Eds.), Discourse comprehension: Essays in honor of Walter Kintsch (pp. 35-48). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ferstl, E. C., & Kintsch, W. (1999). Learning from text: Structural knowledge assessment in the study of discourse comprehension. In H. van Oostendorp & S.R. Goldman (Eds.), The Construction of mental representations during reading (pp. 247-277). Hillsdale, NJ: Lawrence Erlbaum Associates.

Fincham, F.D. (1983). Developmental dimensions of attribution theory. In J. Jaspars, F.D. Fincham, & M. Hewstone (Eds.), Attribution Theory and Research: Conceptual developmental and social dimensions (pp. 117-164). Toronto, Ont: Academic Press.

Fletcher, C.R., van den Broek, P., & Arthur, E.J. (1996). A model of narrative comprehension and recall. In B.K. Britton & A.C. Graesser (Eds.), Models of understanding text (pp. 141-163). Hillsdale, NJ: Lawrence Erlbaum Associates.

Garner, R. (1990). When children and adults do not use learning strategies: Toward a theory of settings. Review of Educational Research, 60, 517-529.

Garner, R., Alexander, P.A., & Hare, V.C. (1991). Reading comprehension failure in children. In B.Y.L. Wong (Ed.), Learning about learning disabilities (pp. 283-307). New York, NY: Academic Press.

Garner, R., Alexander, P., Slater, W., Chou Hare, V., Smith, R., & Reis, R. (1986). Children's knowledge of structural properties of expository text. Journal of Educational Psychology, 78, 411-416.

- Gernsbacher, M.A. (1990). Language comprehension as structure building. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gick, M.L., & Holyoak, K.J. (1980). Analogical problem solving. Cognitive Psychology, 12, 306-355.
- Gick, M.L., & Holyoak, K.J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.
- Glenberg, A.M., & Epstein, W. (1985). Calibration of comprehension. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 702-718.
- Glenberg, A.M., & Epstein, W. (1987). Inexpert calibration of comprehension. Memory and Cognition, 15, 84-93.
- Glenberg, A.M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. Journal of Experimental Psychology: General, 116, 119-136.
- Glenberg, A.M., Wilkinson, A.C., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. Memory and Cognition, 10, 597-602.
- Goodman, K.S. (1984). Unity in reading. In A.C. Purves & O. Niles (Eds.), Becoming readers in a complex society. (Eighty-third yearbook, pp. 79-114). Chicago, IL: The National Society for the Study of Education.
- Goldsmith, T.E., Johnson, P.J., & Acton, W.H. (1991). Assessing structural knowledge. Journal of Educational Psychology, 83, 88-96.

Graesser, A.C. (1981). Prose comprehension beyond the word. New York, NY: Springer-Verlag.

Graesser, A.C., Bertus, E.L., & Magliano, J.P. (1995). Inference generation during the comprehension of narrative text. In R.F. Lorch, Jr. & E.J. O'Brien (Eds.), Sources of coherence in reading (pp. 295-320). Hillsdale, NJ: Lawrence Erlbaum Associates.

Graesser, A.C., & Goodman, S.M. (1985). Implicit knowledge, question answering, and the representation of expository text. In B.K. Britton & J.B. Black (Eds.), Understanding expository text (pp. 109-171). Hillsdale, NJ: Lawrence Erlbaum Associates.

Graesser, A.C., & Kreuz, R. J. (1993). A theory of inference generation during text comprehension. Discourse Processes, 16, 145-160.

Graesser, A.C., & McMahan, C.L. (1993). Anomalous information triggers questions when adults solve problems and comprehend stories. Journal of Educational Psychology, 85, 136-151.

Graesser, A.C., Robertson, S.P., & Anderson, P.A. (1981). Incorporating inferences in narrative representations: A study of how and why. Cognitive Psychology, 13, 1-26.

Graesser, A.C., Robertson, S.P., Lovelace, E., & Swinehart, D. (1980). Answers to why-questions expose the organization of story plot and predict recall of actions. Journal of Verbal Learning and Verbal Behavior, 19, 110-119.

Graesser, A.C., Swamer, S.S., Baggett, W.B., & Sell, M.A. (1996). New models of deep comprehension. In B.K. Britton & A.C. Graesser (Eds.), Models of understanding text (pp. 1-32). Hillsdale, NJ: Lawrence Erlbaum Associates.

Graesser, A.C., & Zwaan, R.A. (1995). Inference generation and the construction of situation models. In C. Weaver, III, S. Mannes, & C.R. Fletcher, (Eds.), Discourse comprehension: Essays in honor of Walter Kintsch (pp. 117-139). Hillsdale, NJ: Lawrence Erlbaum Associates.

Greeno, J.G. (1997). On claims that answer the wrong questions. Educational Researcher, 26, 5-17.

Jacobs, J.B., & Paris, S.G. (1987). Children's metacognition about reading: Issues in definition, measurement and instruction. Educational Psychologist, 22, 255-278.

Jacoby, L.L., & Craik, F.I.M. (1979). Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery of initial context. In L.S. Cermak & F.I.M. Craik (Eds.), Levels of processing in human memory (pp. 1-21). Hillsdale, NJ: Lawrence Erlbaum Associates.

Just, M.A., & Carpenter, P.A. (1987). The psychology of reading and language to comprehension. Boston, MA: Allyn & Bacon.

Kamas, E.N., & Reder, L.M. (1995). The role of familiarity in cognitive processing. In R.F. Lorch, Jr. & E.J. O'Brien (Eds.), Sources of coherence in reading (pp.177-202). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Kiewra, K.A. (1988). Cognitive aspects of autonomous note taking: Control processes, learning strategies, and prior knowledge. Educational Psychologist, 23, 39-56.
- Kiewra, K.A. (1989). A review of note-taking: The encoding-storage paradigm and beyond. Educational Psychology Review, 1, 147-172.
- King, A. (1989). Effects of self-questioning training on college students' comprehension of lectures. Contemporary Educational Psychology, 14, 366-381.
- King, A. (1990a). Enhancing peer interaction and learning in the classroom through reciprocal questioning. American Educational Research Journal, 27, 664-687.
- King, A. (1990b). Reciprocal peer-questioning: A strategy for teaching students how to learn from lectures. The Clearing House, 64, 131-135.
- King, A. (1991a). Effects of training in strategic questioning on children's problem-solving performance. Journal of Educational Psychology, 83, 307-317.
- King, A. (1991b). Improving lecture comprehension: Effects of a metacognitive strategy. Applied Cognitive Psychology, 5, 331-346.
- King, A. (1992a). Facilitating elaborative learning through guided student-generated questioning. Educational Psychologist, 27, 111-126.
- King, A. (1992b). Comparison of self-questioning, summarizing, and note taking-review as strategies for learning from lectures. American Educational Research Journal, 29, 303-323.

King, A. (1994a). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. American Educational Research Journal, 31, 338-368.

King, A. (1994b). Autonomy and question asking: The role of personal control in guided student-generated questioning. Learning and Individual Differences, 6, 163-185.

King, A., & Rosenshine, B. (1993). Effects of guided cooperative questioning on children's knowledge construction. Journal of Experimental Education, 6, 127-148.

Kintsch, E. (1990). Macroprocesses and microprocesses in the development of summarization skill. Cognition and Instruction, 7, 161-195.

Kintsch, W. (1985). Text processing: A psychological model. In T. A. van Dijk (Ed.), Handbook of discourse analysis: Vol. 1. Disciplines of discourse (pp. 231-245).

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. Psychological Review, 95, 163-182.

Kintsch, W. (1989). Learning from text. In L. B. Resnick (Ed.), Knowing, learning, and instruction: Essays in honor of Robert Glaser (pp. 25-45). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kintsch, W. (1992). A cognitive architecture for comprehension. In H.L. Pick, P. van den Broek, & D.C. Knill (Eds.), Cognition: Conceptual and methodological issues (pp. 143-163). Washington, DC: American Psychological Association.

- Kintsch, W. (1993). Information accretion and reduction in text processing: Inferences. Discourse Processes, 16, 193-202.
- Kintsch, W. (1994a). Text comprehension, memory, and learning. American Psychologist, 49, 294-303.
- Kintsch, W. (1994b). Discourse processing. In G. d'Ydewalle, P. Eelen, & P. Bertelson (Eds.), International perspective in psychological science: Volume 2. State of the art (pp. 135-155). Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.
- Kintsch, W. (1994c). The psychology of discourse processing. In M.A. Gernsbacher (Ed.), Handbook of psycholinguistics (pp. 721-739). San Diego, CA: Academic Press.
- Kintsch, W. (1995). How readers construct situation models for stories: The role of syntactic cues and causal inferences. In M.A. Gernsbacher & T. Givon (Eds.), Coherence in Spontaneous Text (pp. 139-160). Philadelphia, PA: John Benjamins Publishing Company.
- Kintsch, W. (1998). Comprehension: A paradigm for cognition. New York, NY: Cambridge University Press.
- Kintsch, W., & van Dijk, T.A. (1978). Towards a model of text comprehension and production. Psychological Review, 85, 363- 394.
- Kletzien, S.B. (1991). Strategy use by good and poor comprehenders reading expository text of differing levels. Reading Research Quarterly, 26, 67-86.

Kluwe, R.H. (1993). Knowledge and performance in complex problem solving. In G. Strube & K.F. Wender (Eds.), The cognitive psychology of knowledge, (pp. 401-425). New York, NY: North-Holland.

Kotovsky, K., & Fallside, D. (1989). Representation and transfer in problem solving. In D. Klahr & K. Kotovsky (Eds.), Complex information processing: The impact of Herbert A. Simon (pp. 69-108). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kourilsky, M., & Wittrock, M.C. (1992). Generative teaching: An enhancement strategy for the learning of economics in cooperative groups. American Educational Research Journal, 29, 861-876.

Kraker, M.J. (1993). Learning to write: Children's use of notation. Reading Research and Instruction: The Journal of the College Reading Association, 32, 55-75.

Kuhn, D. (1992). Cognitive development. In M.H. Bornstein & M.E. Lamb (Eds.), Developmental psychology: An advanced textbook (pp. 211-272). Hillsdale, NJ: Lawrence Erlbaum Associates.

Laidlaw, E.N., Skok, R.L., & McLaughlin, T.F. (1993). The effects of note taking and self-questioning on quiz performance. Science Education, 77, 75-82.

Lawson, M. J., & Chinnappan, M. (1994). Generative activity during geometry problem solving: Comparison of the performance of high-achieving and low-achieving high school students. Cognition and Instruction, 12, 61-93.

Lockhart, R.S., Craik, F.I.M., and Jacoby, L.L. (1976). Depth of processing, recognition, and recall. In J. Brown (Ed.), Recall and recognition (pp. 75-102). London, UK: Wiley.

Lorch, R.F. (1995). Integration of topic information during reading. In R.F. Lorch, Jr. & E.J. O'Brien (Eds.), Sources of coherence in reading (pp.279-294). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lorch, R.F., & Lorch, E.P. (1996). Effects of organizational signals on free recall on expository text. Journal of Educational Psychology, 88, 38-48.

Lorch, R.F., Lorch, E.P., & Inman, W.E. (1993). Effects of signalling topic structure on text recall. Journal of Educational Psychology, 85, 281-290.

Lorch, R.F., Klusewitz, M.A., & Lorch, E.P. (1995). Distinctions among reading situations. In R.F. Lorch & E.J. O'Brien (Eds.), Sources of coherence in reading (pp. 375-398). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lorch, R.F., Jr., & van den Broek, P. (1997). Understanding reading comprehension: Current and future contributions of cognitive science. Contemporary Educational Psychology, 22, 213-246.

Luria, A. (1978). The development of writing in the child. In M. Cole (Ed.), The selected writings of Luria (pp. 145-194) White Plains, NY: M.E. Sharpe.

Lysynchuk, L.M., Pressley, M., & Vye, N.J. (1990). Reciprocal teaching improves standardized reading comprehension in poor comprehenders. The Elementary School Journal, 90, 469-484.

Machiels-Bongaerts, M., Schmidt, H.G., & Boshuizen, H. P.A. (1995). The effect of prior knowledge activation on text recall: An investigation of two conflicting hypotheses. British Journal of Educational Psychology, 65, 409-423.

Maki, R.H., & Berry, S.L. (1984). Metacomprehension of text material. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 663-679.

Maki, R.H., Foley, J.M., Kajer, W.K., Thompson, R.C., & Willert, M.G. (1990). Increased processing enhances calibration of comprehension. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 609-616.

Maki, R.H., & Swett, S. (1987). Metamemory for narrative text. Memory and Cognition, 15, 72-83.

Mandler, G. (1985). Cognitive psychology: An essay in cognitive science. Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.

Mandler, G. (1989). Memory: Conscious and unconscious. In P.R. Solomon, G.R. Goethals, C.M. Kelley, & B.R. Stephens (Eds.), Memory: Interdisciplinary approaches (pp. 84-106). New York, NY: Springer-Verlag.

Mannes, S. M., & Kintsch, W. (1987). Knowledge organization and text organization. Cognition and Instruction, 4, 91-115.

Mannes, S., & St. George, M. (1996). Effects of prior knowledge on text comprehension: A simply modeling approach. In B.K. Britton & A.C. Graesser (Eds.), Models of understanding text (pp. 115-139). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mayer, R.E. (1989). Models for understanding. Review of Educational Research, 59, 43-64.

Mayer, R.E. (1992). Cognition and instruction: Their historic meeting within educational psychology. Journal of Educational Psychology, 84, 405-412.

McClelland, J.L., & Rumelhart, D.E. (1986). A distributed model of human learning and memory. In J.L. McClelland, D.E. Rumelhart, & the P.D.P. Research Group (Eds.), Parallel distributed processing: Vol. 2. Psychological and biological models (pp. 170-215). Cambridge, MA: The MIT Press.

McDaniel, M.A. (1984). The role of elaborative and schema processes in story memory. Memory and Cognition, 12, 46-51.

McDaniel, M.A., & Donnelly, C.M. (1996). Learning with analogy and elaborative interrogation. Journal of Educational Psychology, 88, 508-519.

McDaniel, M.A., Einstein, G.O., Dunay, P.K., & Cobb, R.E. (1986). Encoding difficulty and memory: Toward a unifying theory. Journal of Memory and Language, 25, 645-656.

McDaniel, M.A., Ryan, E.B., & Cunningham, C.J. (1989). Encoding difficulty and memory enhancement for young and older readers. Psychology and Aging, 4, 333-338.

McGivern, J.E., Levin, J.R., Pressley, M., & Ghatala, E. (1990). A developmental study of memory monitoring and strategy selection. Contemporary Educational Psychology, 15, 103-115.

McKeown, M.G., Beck, I. L., Sinatra, G.M., & Loxterman, J.A. (1992). The contribution of prior knowledge and coherent text to comprehension. International Reading Association, 27, 79-93.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. Psychological Review, 99, 440-466.

McKoon, G., & Ratcliff, R. (1995). The minimalist hypothesis: Directions for research. In C. Weaver, III, S. Mannes, & C.R. Fletcher, (Eds.), Discourse comprehension: Essays in honor of Walter Kintsch (pp. 97-116). Hillsdale, NJ: Lawrence Erlbaum Associates.

McNamara, D.W., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. Discourse Processes, 22, 247-288.

McNamara, D.S., Kintsch, E., Songer, N.B., & Kintsch, W. (1996). Are good texts always better? Text coherence, background knowledge, and levels of understanding in learning from text. Cognition and Instruction, 14, 1-43.

Messick, S. (1984). The psychology of educational measurement. Journal of Educational Measurement, 21, 215-237.

Meyer, B.J.F. (1985). Prose analysis: Purposes, procedures, and problems. In B.K. Britton & J.B. Black (Eds.), Understanding expository text (pp. 11-64). Hillsdale, NJ: Lawrence Erlbaum Associates.

Moravcsik, J.E., & Kintsch, W. (1993). Writing quality, reading skills, and domain knowledge as factors in text comprehension. Canadian Journal of Experimental Psychology, 47, 360-374.

Morris, C.C. (1990). Retrieval processes underlying confidence in comprehension judgments. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 223-232.

Moshman, D. (1982). Exogenous, endogenous, and dialectical constructivism. Developmental Review, 2, 371-384.

Murdock, B.B. (1995). Primacy and recency in the chunking model. In C. Weaver, III, S. Mannes, & C.R. Fletcher, (Eds.), Discourse comprehension: Essays in honor of Walter Kintsch (pp. 49-63). Hillsdale, NJ: Lawrence Erlbaum Associates.

Nelson, T.O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. Psychological Bulletin, 95, 109-133.

Nicholls, J.G. (1978). The development of the concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. Child Development, 49, 800-814.

Nicholls, J.G. (1983). Conceptions of ability and achievement motivation: A theory and its implications for education. In S.Paris, G. Olson, & H. Stevenson (Eds.), Learning and motivation in the classroom (pp. 211-237). Hillsdale, NJ: Lawrence Erlbaum Associates.

Nicholls, J.G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. Psychological Review, 19, 308-346.

Nickerson, R. S. (1989). New directions in educational assessment. Educational Researcher, 18, 3-7.

Norris, S. P. (1989). Can we test validly for critical thinking? Educational Researcher, 18, 21-26.

O' Brien, E.J. (1995). Automatic components of discourse comprehension. In R.F. Lorch & E.J. O'Brien (Eds.), Sources of coherence in reading (pp. 159-175). Hillsdale, NJ: Lawrence Erlbaum Associates.

Olson, D.R. (1994). The world on paper: The conceptual implications of writing and reading. Cambridge, UK: Cambridge University Press.

Olson, D.R. (1996). Literate mentalities: Literacy, consciousness of language, and modes of thought. In D.R. Olson & N. Torrance (Eds.), Modes of thought: Explorations in culture and cognition (pp. 141-151). Cambridge, UK: Cambridge University Press.

Otero, J., & Kintsch, W. (1992). Failures to detect contradictions in a text: What readers believe versus what they read. Psychological Science 3, 229-235.

Palincsar, A.S., & Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. Cognition and Instruction, 1, 117-175.

Paris, S.G., & Newman, R.S. (1990). Developmental aspects of self-regulated learning. Educational Psychologist, 25, 87-102.

Paris, S.G., & Winograd, P. (1990). How metacognition can promote academic learning and instruction. In B.L. Jones & L. Idol (Eds.), Dimensions of thinking, and cognitive instruction (pp. 15-51). Hillsdale, NJ: Lawrence Erlbaum Associates.

Paul, R.W. (1990). Critical and reflective thinking: A philosophical perspective. In B.F. Jones & L. Idol (Eds.), Dimensions of thinking and cognitive instruction (pp. 445-494). Hillsdale, NJ: Lawrence Erlbaum Associates.

Peper, R.J., & Mayer, R. E. (1986). Generative effects of note-taking during science lectures. Journal of Educational Psychology, 78, 34-38.

Perfetti, C.A., & Britt, M. A. (1995). Where do propositions come from? In C. Weaver, III, S. Mannes, & C.R. Fletcher, (Eds.), Discourse comprehension: Essays in honor of Walter Kintsch (pp. 11-34). Hillsdale, NJ: Lawrence Erlbaum Associates.

Piaget, J. (1970). The science of education and the psychology of the child. New York, NY: Orion Press.

Pressley, M. (1997). Commentary: The cognitive science of reading. Contemporary Educational Psychology, 22, 247-259.

Pressley, M., & Dennis-Rounds, J. (1980). Transfer of a mnemonic keyword strategy at two age levels. Journal of Educational Psychology, 72, 575-582.

Pressley, M., Harris, K.R., & Marks, M.B. (1992). But good strategy instructors are constructivists! Educational Psychology Review, 4, 3-31.

Pressley, M., McDaniel, M.A., Turnure, J.E., Wood, E., & Ahmad, M. (1987). Generation and precision of elaboration: Effects on intentional and incidental learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 291-300.

Pressley, M., Symons, S., McDaniel, M.A., Snyder, B.L., & Turnure, J.E. (1988). Elaborative interrogation facilitates acquisition of confusing facts. Journal of Educational Psychology, 80, 268-278.

Putz-Osterloh, W. (1993). Strategies for knowledge acquisition and transfer of knowledge in dynamic tasks. In G. Strube & K.F. Wender (Eds.), The cognitive psychology of knowledge, (pp. 331-351). New York, NY: North-Holland.

Reder, L.M., & Kusbit, G.W. (1991). Locus of the Moses Illusion: Imperfect encoding, retrieval, or match? Journal of Memory and Language, 30, 385-406.

Reynolds, R.E., Sinatra, G.M., & Jetton, T.L. (1996). Views of knowledge acquisition and representation: A continuum from experience centered to mind centered. Educational Psychologist, 31, 93-104.

Risch, N. L., & Kiewra, K.A. (1990). Content and form variations in note taking: Effects among junior high students. Journal of Educational Research, 83, 355-357.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. Review of Educational Research, 66, 181-221.

Rumelhart, D.E., Hinton, G.E., & McClelland, J.L. (1986). A general framework for parallel distributed processing. In D.E. Rumelhart, J.L. McClelland, & the PDP Research Group (Eds.), Parallel distributed processing: Exploration in the microstructure of cognition: Vol. 1. Foundations (pp. 45-77). Cambridge, MA: The MIT Press.

Rumelhart, D.E., McClelland, J.L., & the PDP Research Group (1986). Parallel distributed processing: Exploration in the microstructure of cognition: Vol. 1. Foundations. Cambridge, MA: The MIT Press.

Ryan, F.L. (1971). Exemplars for the new social studies: Instructing in the elementary school. New Jersey: Prentice-Hall.

Ryan, M.P. (1984). Monitoring text comprehension: Individual differences in epistemological standards. Journal of Educational Psychology, 76, 248-258.

Sattler, J.M. (1992). Assessment of children (Rev. and Updated 3rd ed.). San Diego, CA: Jerome M. Sattler, Publisher, Inc.

Scardamalia, M., & Bereiter, C. (1984). Development of strategies in text processing. In H. Mandl, N.L. Stein, & T. Trabasso (Eds.), Learning and comprehension of text (pp. 379-406). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Schank, Roger C. (1986). Explanation patterns: Understanding mechanically and creatively. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schmalhofer, F. (1995). The acquisition of knowledge from text and example situations: An extension to the construction-integration model. In C. Weaver, III, S. Mannes, & C.R. Fletcher, (Eds.), Discourse comprehension: Essays in honor of Walter Kintsch (pp. 257-283). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schmidt, H.G., De Volder, M.L, De Grave, W.S., Moust, J. H. C., & Patel, V. L. (1989). Explanatory models in the processing of science text: The role of prior knowledge activation through small-group discussion. Journal of Educational Psychology, 81, 610-619.

Schunk, D.H. (1994). Self-regulation of self-efficacy and attributions in academic settings. In D.H. Schunk & B.J. Zimmerman (Eds.), Self-regulation of learning and performance (pp. 75-101). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schunk, D.H., & Rice, J.H. (1987). Enhancing comprehension skill and self-efficacy with strategy value information. Journal of Reading Behavior, 3, 285-302.

Sharkey, N.E. (1990). A connectionist model of text comprehension. In D.A. Balota, G.B. Flores d'Arcais, & K. Rayner (Eds.), Comprehension processes in reading (pp. 487-514). Hillsdale, NJ: Lawrence Erlbaum Associates.

Spires, H. A. (1993). Learning from a lecture: Effects of comprehension monitoring. Reading Research and Instruction, 32, 19-30.

Spires, H.A., & Donley, J. (1998). Prior knowledge activation: Inducing engagement with informational texts. Journal of Educational Psychology, 90, 249-260.

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. Reading Research Quarterly, 20, 360-406.

Stanovich, K.E. (1989). Has the learning disabilities field lost its intelligence? Journal of Learning Disabilities, 22, 487-492.

Stein, N.L., Bernas, R.S., Calicchia, D.J., & Wright, A. (1996). Understanding and resolving arguments: The dynamics of negotiation. In B.K. Britton & A.C. Graesser (Eds.), Models of understanding text (pp. 257-287). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R.J. (1985). Beyond I.Q.: A triarchic theory of human intelligence. Cambridge, UK: Cambridge University Press.

Sternberg, R.J. (1988). The triarchic mind: Conceptions of the nature of intelligence. Cambridge, UK: Cambridge University Press.

Tapiero, I., & Denhiere, G. (1995). Simulating recall and recognition by using Kintsch's construction-integration model. In C. Weaver, III, S. Mannes, & C.R. Fletcher, (Eds.), Discourse comprehension: Essays in honor of Walter Kintsch (pp. 211-232). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tobias, S. (1994). Interest, prior knowledge, and learning. Review of Educational Research, 64, 37-54.

van den Broek, P., Fletcher, C.R., & Risdén, K. (1993). Investigations of inferential processes in reading: A theoretical and methodological integration. Discourse Processes, 16, 169-180.

van den Broek, P., Risdén, K., Fletcher, C.R., & Thurlow, R. (1996). A "landscape" view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In B.K. Britton & A.C. Graesser (Eds.), Models of understanding text (pp. 165-187). Hillsdale, NJ: Lawrence Erlbaum Associates.

van den Broek, P., Risdén, K., & Husebye-Hartmann, E. (1995). The roles of readers' standards for coherence in the generation of inferences during reading. In R.F. Lorch & E.J. O'Brien (Eds.), Sources of coherence in reading (pp. 353-176). Hillsdale, NJ: Lawrence Erlbaum Associates.

van Dijk, T.A. (1995). On macrostructures, mental models, and other inventions: A brief personal history of the Kintsch-van Dijk theory. In C. Weaver, III, S. Mannes, & C.R. Fletcher, (Eds.), Discourse comprehension: Essays in honor of Walter Kintsch (pp. 383-410). Hillsdale, NJ: Lawrence Erlbaum Associates.

van Dijk, T.A., & Kintsch, W. (1983). Strategies of discourse comprehension. San Diego, CA: Academic Press.

von Glaserfeld, E. (1984). An introduction to radical constructivism. In P. Watzlawick (Ed.), The invented reality (pp. 17-40). New York, NY: Norton.

Voss, J.F., & Bisanz, G.L. (1985). Knowledge and the processing of narrative and expository texts. In B.K. Britton & J.B. Black (Eds.), Understanding expository text (pp. 173-226). Hillsdale, NJ: Lawrence Erlbaum Associates.

Vygotsky, L. (1978). Mind in society: The development of higher psychological processes (M. Cole, V. John-Steiner, S. Scribner & E. Soubberman, Eds. and Trans.). Cambridge, MA: Harvard University Press.

Vygotsky, L. (1986). Thought and language (A. Kozulin, Ed. and Trans.). Cambridge, MA: The MIT Press.

Vygotsky, L. (1987). The collected works of L.S. Vygotsky: Vol. 1. Problems of General Psychology (R.W. Reiber & A.S. Carton, Eds. and Trans.). New York, NY: Plenum Press.

Wandersee, J.H. (1988). Ways students read texts. Journal of Research in Science Teaching, 25, 69-84.

Weaver, C.A. (1990). Constraining factors in calibration of comprehension. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 214-222.

Weaver, C.A., Bryant, D.S., & Burns, K.D. (1995). Comprehension monitoring: Extensions of the Kintsch and van Dijk model. In C. Weaver, III, S. Mannes, & C.R. Fletcher, (Eds.), Discourse comprehension: Essays in honor of Walter Kintsch (pp. 177-193). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Wittrock, M.C. (1990). Generative processes of comprehension. Educational Psychologist, 24, 345-376.
- Woloshyn, V.E., Paivio, A., & Pressley, M. (1994). Use of elaborative interrogation to help students acquire information consistent with prior knowledge and information inconsistent with prior knowledge. Journal of Educational Psychology, 86, 79-89.
- Wong, B.Y.L. (1985). Self-questioning instructional research: A review. Review of Educational Research, 55, 227-268.
- Wood, E., Pressley, M., & Winne, P.H. (1990). Elaborative interrogation effects on children's learning of factual content. Journal of Educational Psychology, 82, 741-748.
- Zwaan, R.A. (1996). Toward a model of literary comprehension. In B.K. Britton & A.C. Graesser (Eds.), Models of understanding text (pp. 241-255). Hillsdale, NJ: Lawrence Erlbaum Associates.

APPENDIX A

Constructivist Question Stems

Constructivist Question Stems*

	<i>Models</i>
1. What do you think might occur if.....? answer:.....	situation
2. What information do we already have about.....? How does it apply to? answer:.....	textbase/situation
3. Are there any differences between.....and....? Explain	textbase/situation
4.appears to be a problem because..... What are some possible solutions?	textbase/ situation
5. The author(s) states that "....." Explain why this statement is true or false.	textbase/situation
6. Compare.....and/within regards to..... Explain your answer.	textbase/situation

While textbase questions contain information which is stated in a single sentence of the text, situation questions require inferences. Therefore, listed below are categories of situation model inferences which may be created by the students. Please note that they are preceded by question stem numbers which lend themselves to potential inference categories.

Situation model Inferences

#2., #3., #4., & #6. *bridging inference* - "requires linking information from two or more sentences in the text to answer the question. Inferring the unstated relation between sentences is also a process that relies on the situation model" (McNamara et al., 1996).

-Two or more sentences or concepts have to be linked to make and answer the question.

#1., #2., #3, #4., #5., & #6= *elaborative inference* - "requires linking textual and outside knowledge information, which requires some, but not necessarily a very deep, situation understanding."

-Use of knowledge about the topic to fill-in additional detail not mentioned in the text or to establish connections between what is being read and related items of knowledge.

#1. & #4. - *problem solving* - "requires applying information from the text to a novel situation and hence depend on situational understanding"

Coherence, whether it be local or global, is increased by the strategies listed below. Cited are particular questions stems which may incite the building of greater coherence.

Coherence strategies

#2.,#3.,#4., & #6. = *bridging inference*

#2.,#3.,#4., & #6 = *argument repetition* - has referential ties and a common referent.

- "two propositions are related if they share a common argument. Coherence is thus reduced to referential ties, which is an oversimplification, but attractively simple. - may often be merely an accidental by-product of some more basic coherence relationships among propositions but serves to index the existence of a relationship in an objective, easily identifiable way".

- Is predicative of recall (van Dijk & Kintsch, 1983).

#1., #3.,#4., & #6. = *causal connection* - relations among events and actions.

- "exists between states and events in the physical world" (van Dijk & Kintsch, 1983).

*Adapted from Ryan (1971) and King (e.g., 1989)

APPENDIX B

Domain Knowledge Questionnaire

Knowledge Questionnaire

Title: The Effects of Question Construction on Expository Text Comprehension

Researcher: Thelma M. Gunn

Supervisor: Dr. B.S. Randhawa

Professor, Department of Educational Psychology

A.Preliminary Information. Answer the following questions as honestly and as fully as possible.

1. Before graduating from high school, did you take any courses relating to human anatomy?

Yes

No

If "yes", briefly identify when you were enrolled and those aspects of human anatomy that you studied.

2. Have you ever taken any classes pertaining to human anatomy while in university?

Yes

No

If "yes", briefly identify when you were enrolled and those aspects of human anatomy that you studied.

3. Have you ever taken any emergency medicine courses (i.e., C.P.R., E.M.T., etc.)?

Yes

No

If "yes", specify the name of the course, when you were enrolled, and what it entailed.

4. Have you ever had a close personal relationship with someone with heart problems or heart disease?

Yes

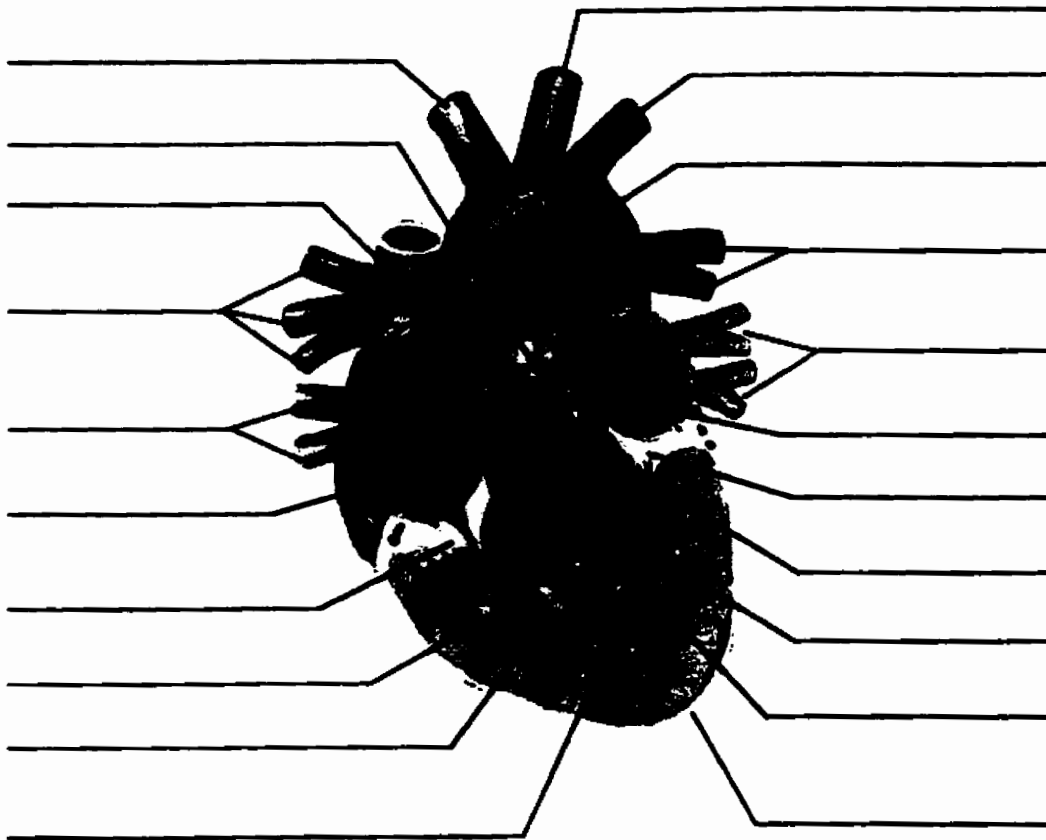
No

If "yes", specify the nature of your relationship (i.e., immediate family, close personal friend, etc.) and the nature of the heart problem or disease.

B. Diagram Labelling.

5. This is a diagram of a human heart.

a). In the spaces provided, name as many parts of the heart as you can. If you are uncertain of the names and locations of certain parts, you are encouraged to guess.



C. Multiple Choice. There is only one answer that is either correct or most appropriate. Circle the answer that corresponds to the question.

6. Which of the following chambers pumps blood into the systemic circulatory system?

- a. right atrium
- b. left atrium
- c. right ventricle
- d. left ventricle

7. Venous (deoxygenated) blood is received by the

- a. right atrium
- b. left atrium
- c. right ventricle
- d. left ventricle

8. The membranous sac surrounding the heart is known as the

- a. pericardium
- b. myocardium
- c. endocardium
- d. epicardium

9. A myocardial infarction results in

- a. the death of an area of the aorta
- b. an accelerated rate of hemopoiesis
- c. rapid cell division of the layers of the pericardium
- d. death of an area of the heart muscle

10. The blood flows from the lungs into the heart's

- a. right atrium
- b. left ventricle
- c. atrioventricular valve
- d. left atrium

11. A "blue baby" is probably

- a. suffering from rheumatic heart disease
- b. suffering from arterial sclerosis
- c. not getting enough oxygenated blood throughout the body
- d. not getting enough iron in the body.

12. The heart beat originates in the

- a. AV node
- b. pacemaker
- c. automatic nervous system
- d. pericardium

13. What is the normal heart beat for a young adult?

- a. 100 beats per minute
- b. 90 beats per minute
- c. 70 beats per minute
- d. 50 beats per minute

14. Coronary disease is the result of
- a. high blood pressure
 - b. a poorly shaped coronary valve
 - c. a bacterial infection to the myocardium
 - d. blocked coronary arteries
15. A septal defect can be described as
- a. a misshaped heart valve
 - b. a gap between the two sides of the heart
 - c. arterial walls which are lumpy, rough, and narrow in shape
 - d. a thinning of the lining of the endocardium

Purpose for Study: This is a two phase study. The purpose of the first phase of the study has been to determine those individuals who have high knowledge and those who have low knowledge of the heart and its functions. You have completed the first phase of the study.

The purpose of the second phase of the study is to examine whether a specific reading comprehension technique (i.e., questioning) will not only enable students with high domain knowledge and low domain knowledge to better comprehend expository text, but help them to monitor their reading behaviours. Students who are eligible (i.e., have high or low domain knowledge of the heart) and who are willing to participate will be randomly assigned to three strategic conditions: highly structured, semi-structured, and unstructured questions. Following a brief training period, they will be asked to read a few pages of scientific text. Some informal testing will occur immediately afterward. This will include a brief comprehension monitoring questionnaire, a 12-item posttest, a recall task, and a word association task. All instruction and testing will be conducted in groups of 10 to 20 participants.

Duration of Study: 120 minutes

Credits: 2 credits for the first hour and \$5 for the second hour

Location, Dates and times: To be announced

If you are eligible to participate in the second phase of the study, would you be interested?

YES

NO

If "yes", please provide your name and telephone number so that you may be easily contacted. Thank you for your participation thus far.

Name _____
(print name)

Telephone Number _____

RESEARCHER'S INFORMATION

----- SCORING

- section 1 - 1 point if "yes"
 - 1 point if they took the class in the last 2 years, .5 if they took the class more than 2 years ago
 - 1 point if they name the heart/circulation
 - question #4, 1 point if "yes" and 1 point if can specifically name the nature of the disease, .5 point if give a general response

(maximum 11 points)

- section 2 - 2 points for each category answered in full, 1 point if partial answer in correct category, .5 point if give a correct name but in wrong spot, .5 point if they define the function but cannot come up with the name.

(maximum 42 points)

- section 3 - 2 points if correct
- (maximum 20 points)

Maximum total points = 73

Section 2 Answers (Clockwise):

Left common carotid artery; left subclavian artery; aorta; left pulmonary arteries; left atrium; mitral valve; aortic valve; infundibulum; left ventricle; apex; interventricular septum; trabeculations; right ventricle; tricuspid valve; right atrium; right pulmonary veins; right pulmonary arteries; superior vena cava; pulmonic valve; brachiocephalic artery.

Section 3 Answers:

6) d; 7) a; 8) a; 9) d; 10) d; 11) c; 12) b; 13) c 14) d; 15) b

APPENDIX C
Training Materials

PROCEDURE

Screening:

1. **Prior knowledge questionnaire** (10 minutes)

Experimental session:

1. **Practice session** (20 minutes)
 - generic question stem and signal word
 - i. question training regarding memory (textbase model) and critical thinking (situation model) questions.
 - ii. strategy prompt cards with practice text
 - unguided questioning
 - no question training or practice
2. **Pre-reading sorting task** (10 minutes)
3. **Reading task** (10 minutes)
 - Text entitled "Blood" (5 minutes) for all subjects.
 - Followed by the experiment text entitled "Heart Disease" (5 minutes).
 - *instruct students to reread each text if there is time remaining
3. **Experimental questioning tasks** (20 minutes)
 - Students will begin individual question generation using strategy prompt cards and may use all texts affiliated with their experimental condition.
 - Record questions and answers on question/response sheets
4. **Calibration of comprehension questionnaire** (5 minutes)
 - *instruct students to check their answers if there is time remaining
5. **Summary recall text** (10 minutes)
 - *instruct students to check their answers if there is time remaining
6. **Post test** (12 questions - 30 minutes)
 - *instruct students to check their answers if there is time remaining
7. **Post-reading sorting task** (10 minutes)
 - *instruct students to check their answers if there is time remaining

115 minutes

Questioning training

1. Before we start I will briefly tell you what we will be doing during this session. (Procedure). Everything is timed so I will typically give you a 2 or 5 minute warning; depending on the length of the task.
2. I have given you a consent form to sign before we start. Please read it over now and sign it. (please print) Leave it at the top of the desk in front of you. I will pick it up in a few minutes. At the end of the experiment I will give you a copy of the consent form, a review of the experiment, and your credits/money.
3. I will be giving you identification numbers. There is a slot at the top of all tests which requires this number. Please insert it whenever it is required. Whenever you are finished with an exam or any materials please leave them face down in front of you and I will pick them up. If you have any new materials in front of you and you are finished early, you are not allowed to read or look at them until the next task begins.

Introduction

-As stated earlier, the purpose of this experiment is to understand how people read and comprehend text.

-I am going to start by saying that reading and comprehension processes have typically been difficult to understand. This is especially true in the case of expository text. Unlike narrative text, expository textbooks typically give you a lot of new information and they are written in a way that can be difficult to understand or follow. However, there are various ways that may help you to understand what you are reading. Some of you have developed your own ways or strategies for understanding and others have been taught some useful strategies. One comprehension technique that you may be familiar with is "questioning". This can be done in several ways. For instance, some people ask themselves questions as they are reading and some ask themselves questions after they have finished reading. There are also different types of questions which can be asked: questions about what was written in the text and those that require you to think about the information in more critical ways. Therefore, questioning can help you to understand and comprehend material that is new or complex. That is, what we are attempting to understand with your help.

Signal word and generic question stems only

I am now going to explain a bit more about the two types of questions that I just mentioned. That is, there are questions that ask for information straight out of the text (i.e., memory questions) and there are questions that require you to think more critically (i.e., critical thinking questions). To demonstrate these types of questions, I am going to begin by asking you to quickly read the text entitled "Indonesia and South Korea". (read) (2 minutes)

The first type of question to be discussed is the **memory question**. (overhead) **Memory** questions simply ask you to recall information that you have read. Information that you can simply find by looking back at a specific part of the text or at a specific sentence.

1. For example in this text I could ask "Describe what Indonesia is like geographically"? - the answer can be explicitly found in paragraph 2. A: it consists of thousands of islands, many of them tiny, some huge, which are scattered widely over the balmy, southern seas. There are palm trees along the white sand beaches and the interior is covered by green jungles.

2. "What are the primary products made in Korea today"? Again the answer can be found quickly in paragraph 4. A: textiles such as shirts and dresses and economy cars.

3. "Where do many Indonesians live"? This answer can be found in the first sentence of paragraph 6. A: in the slums of the big city.

Therefore, the questions and answers are easily available within the text. They are recorded facts and they can be memorized. You are not required to think about the material, just provide the easily located answer.

The other type of questions that I mentioned are **critical thinking questions**. These types of questions require you to explain concepts or relationships between ideas, apply the information you have read to other situations, to draw conclusions based on what you have read, to support your ideas based on what you have read, and so forth. Therefore, critical thinking questions ask for more than what you just read. The answer will not simply be in the text or in a sentence. You have to think about what you read and apply that to other parts of the text or what you already know about a topic.

1. For example, "Which country is farther north and higher in elevation" A: You will have to think about what you have read. It mentions in paragraph 3 that Korea has barren mountain ranges as well as cold winds from the north - they also have winter - On the other hand, in paragraph 2 it says that Indonesia consists of islands in the south seas, hot weather, muggy temperatures. Therefore, with some thought, you can infer that Korea is farther north and higher in elevation.

2. "What industries do you think Indonesia bases much of its economy on"? In paragraph 2 it says that the beautiful geography greets many tourists, that there is rice farming (paragraph 5) and we can surmise that because of the weather, there tropically grown food products.

3. "Based on what you have read about Indonesia and Korea, which country has a better standard of living? Support your response" It does not directly state in the text that Korea has a better standard of living but we can surmise that it does because Korea - has more factories, live close to factory employment, no need for child care services, people go to university or technical schools, children can study abroad. Compared to Indonesia - live in

slums, little work, poor education system due to ethnic diversity, poor communication because of the diversity in languages, and their economy is reliant upon weather and the global economy.

Therefore, you can see that critical thinking questions ask you to “think” , not just remember or repeat what you have read.

*While memory questions have merit, they aren't as helpful as critical thinking questions when it comes to remembering and understanding text material. Therefore, you want to try and develop critical thinking questions as much as possible. However, it isn't always easy to think of good questions. Sometimes it is convenient to have a prompt or a cue for generating the best questions. The next part of our discussion is going to involve ways to make good questions using a particular strategy.

A. Signal Words

I am sure that most of you have heard of the the 5Ws = “who, what, why, when, where, and how”. These are considered to be fundamental questioning tools used by journalists. These words signal the questioner to acquire all of the most important information. Signal words also help the questioner generate memory and critical thinking questions. As mentioned before, critical thinking questions are more in-depth and help the questioner understand and remember the material better. Let's use an example. I am going to ask you to read the text entitled “traits of mammals”. When you are finished we will go through some questions. (3 minutes)

Here are some memory questions using signal words:

1. Who is included in the list as being a mammal? (a dog, monkey & whale, p.1)
2. What do mammals eat? (many different kinds of food, p.5)
3. Where do baby mammals develop before being born? (inside the mother, p.2)
4. When do mammals hibernate? (during winter, p.9)
5. Why do mammals have hair? (as an insulation against heat and cold, p.4)
6. How do reptiles stay warm? (by using the sun, p.3)

Conversely, here are some critical thinking questions:

1. Who is the parent of primary importance for a young mammal? Explain. (mothers - as a food source (i.e., milk) and as a teacher (i.e., hunting))
2. What other species not listed in the text are classified as mammals? Explain. (humans, domestic farm animals, etc because they have the same characteristics listed)

3. Where are mammals found? How are they able to live in these environments? (in many environments because they have developed special ways of protecting their young against danger)
4. When do mammals learn their specialized behaviours? (they are either taught by their mother when young or they are inherent, complex behaviours)
5. Why does a meat-eating mammal have different teeth as compared to a plant-eating mammal? (so that it can cut meat as compared to grinding plant material)
6. How do mammals survive winter when hibernating? (they maintain a low body temperature which requires little energy)

Therefore, you can see that with signal words you can create memory and critical thinking questions. That is, you can create questions which require information simply from the text (i.e., memory) and questions which require you to link ideas and to think about what was written (i.e., critical thinking). The object is to acquire as much information as possible so that you have a clear understanding and memory of what was read.

I am going to ask you to read the text entitled "Peru and Argentina" and then we will try to generate some questions using the signal words. Try to develop critical thinking questions as often as possible (2 minutes).

- ex.
1. Who brought Christianity to Peru?
 2. What would happen to the economy of Argentina if their cattle industry was completely eradicated?
 3. Where do you think Argentineans import cars, machinery, and clothing from?
 4. When was the Incan empire destroyed?
 5. Why is cattle farming such a large industry in Argentina?
 6. How is the geography of Peru similar to the geography of South Korea?

B. Generic question stems

The questioning strategy that I am about to show you involves creating questions using question stems. (Give them examples of the question stems). As you can see, the main part of the question has already been developed. Your responsibility is to think of ways to complete the question by filling-in the missing portions. The question stems frequently ask you to link information to that which you already know or have just read as well as to link information across the various parts of the text. In other words, the stems can be used to

create memory questions, but more often than not they are used to create critical thinking questions. Let's see how they work using a piece of text. I would like you to read the text entitled "Traits of Mammals" and then I will show you some question examples (3 minutes).

1. What do you think might occur if a *mammal was developed outside of its mother's body*? What dangers would there be for the baby's survival?

*2. What information do we already have about *reptiles versus mammals*? How do *these differences* apply to body temperature?

3. Are there any differences between *learned* and *instinctive behavior*? Explain.

4. The *untimely death of a mother* appears to be a problem because of her role as *nurturer and teacher*. What are some possible solutions for the baby's survival if she should die prematurely?

5. The author(s) states that "*Mammals...are the only animals that have hair or fur.*" Explain why this statement is true or false.

6. Compare *incisors and canines* with *premolars and molars* in regards to eating. Explain your answer.

As you can see, many of these questions require you to develop critical thinking questions while others allow for memory questions. The object of the stems is to acquire as much information as possible so that you have a clear understanding and memory of what was read. I am going to ask you to read the text entitled "Peru and Argentina" (2 minutes) and then we will try to generate some questions using the question stems. Remember, some of the stems can be slightly modified to suit your question needs. Also, try to develop critical thinking questions as often as possible.

ex.

1. What do you think might occur to the **economy of Argentina** if their **cattle industry was completely eradicated**?

2. What information do we already have about the **economic potential of cottage industries**? How does it apply to **Peru**?

3. Are there any differences between **Argentinean democracy** and **North American democracy**?

4. Selling beef and wheat appears to be a problem for Argentina because they have difficulty finding buyers. What are some possible solutions?

5. The author states that "In Argentina....there are very few Indians today". explain why this statement is true or false.

6. Compare Peru and South Korea in regards to geography. Explain your answer

*** Explain to the students that these questions require you to link with information that you have already read to information that you already know. Therefore, we are just showing them that if they had information regarding these topics, this is how the information they have just read may apply.**

Instructions:

1. I am now going to ask you to do a category sort. You can put the words into any category you see fit. You can use a word more than once. And there are no correct ways to make categories. Be sure to read the instructions before you start. You have 10 minutes to complete this task. (I will be bringing your numbers around right away so please get started).

2. I am now going to provide you with a text entitled "Blood". Please read the text carefully. You have 5 minutes for this task. If you are done early, please read the text again.

3. I am now going to ask you to read the text entitled "Heart Disease". Again, please read the text carefully. You have 5 minutes for this task. If you are done early, please read the text again.

4. I am now asking you to generate questions concerning the "Heart Disease" text. (You are to use the question stem/signal word strategy). You may refer to and create links with the text entitled "blood", but your chief responsibility is to create questions using the "heart disease" text. You are to create and answer 6 questions on the question/answer sheet provided. (That means that you must use all 6 of the question stems) (That means that you are to use all 6 of the signal words). If you have difficulty with a question, do as much as you can and return to finish it later. You are encouraged to refer to the text as often as possible when creating your questions. You have 20 minutes for this task.

5. This is a confidence measure. It concerns the text entitled "heart disease". After you have read each question, please circle a single number on the accompanying scale to report how confident you are that you could correctly

answer the question being asked. You have 5 minutes to complete this task.

6. You are being asked to summarize what you have read in the text entitled "heart disease" using an essay format. Therefore, you are being asked to state the important information in the text in a briefer form. You have 10 minutes to write all that you can recall.

7. This is a post-test . Answer all of the following questions as carefully and as fully as possible. You may use point form, but be sure that your responses are clear and concise. You have 30 minutes to complete this task.

**8. This last task is the category sort. It is the same sort as you saw at the start of the experiment. Again, you have 10 minutes to complete this task.
(read instructions)**

MEMORY QUESTIONS

- Memory questions ask you to **recall** information that you have read.
- The **information** can be found by simply looking back at a **specific part** of the text or at a **specific sentence**.
- **Questions and answers** are easily available by looking at the text. They are **recorded explicitly** within the text and they can be **memorized**.

Examples:

1. Describe what Indonesia is like geographically?
2. What are the primary products made in Korea today?
3. Where do many Indonesians live?

CRITICAL THINKING QUESTIONS

- Require you to **explain** concepts or relationships between ideas, to **apply** the information that you have read to other situations, to draw **conclusions** based on what you have read, to **support** your ideas based on what you have read, etc.
- Critical thinking questions ask you to **think** about what you have read and to **apply** that information to **other parts of the text** or to what you **already know** about a topic.
- They ask you to critically examine, and not just remember or repeat what you have read.

Examples:

1. Which country is farther north and higher in elevation?
2. What industries do you think Indonesia bases much of its economy on?
3. Based on what you have read about Indonesia and Korea, which country has a better standard of living?

Indonesia and South Korea

Many factors influence whether a developing country can look forward to a prosperous future, or whether it will be forever doomed to a backwardness and a poverty. A comparison between South Korea and Indonesia illustrates this point.

Indonesia consists of thousands of islands, many of them tiny, some huge, which are scattered widely over the balmy, southern seas. Along the coastlines palm trees waving over white sand beaches greet the tourists, who come in their cruise ships. The hot and muggy interior of the islands is covered by green jungles, where many different kinds of plants and animals live. The people are proud and beautiful, and lead carefree lives.

In contrast, the short and stocky Koreans inhabit a peninsula which is criss-crossed by barren mountain ranges. Only along a few rivers and in the plains by the sea is the soil fertile enough to farm. Dry winds blow from the north in winter, and people have to wear heavy, padded jackets against the cold, while in summer the hot sun burns the mountain sides.

Throughout their history Korean farms could barely grow enough cabbage and rice to feed themselves. Now, however, many of them work in textile mills, where they make shirts and dresses which are sold all over the world, because they are good and cheap. Recently, South Korea also began to produce sleek, little economy cars which have been quite a success among American car buyers.

Indonesia differs in many respects. In the old days it was known as the Spice Islands, because merchants from all over the world sailed into its harbors to trade for cinnamon and pepper. Today farms still plant rice on terraced hillsides using the same methods as their forefathers. Each year, more of the jungle is cut down and turned into farmland. As a result the soil becomes exhausted quickly and erodes away when it rains.

Many Indonesians live in the slums of big cities, where they can't find work. Out on the islands the people speak different languages, so that the teachers, whom the government sends to open up schools, often can't talk to their pupils. Most of the people are Muslim, but many of the islands are inhabited by Christians, Hindus, Buddhists, and even primitive tribes who still live in the stone age.

On the other hand, the typical Korean has a job in some factory and lives in a town close by. While the parents work, children go to school, and the grandparents stay at home to take care of the babies. The better students compete for admission to the universities and technical colleges. The whole family is filled with pride when their son gets a scholarship to study in Japan or in the U.S.

Thus, Indonesia and South Korea may develop very differently in the future.

Traits of Mammals

What do a dog, a bat, a monkey, and a whale have in common? They are all mammals. Mammals are a special class of vertebrates (animals with backbones) that share certain traits or characteristics. Mammals are found on land, in water, and in the air because they have developed traits that allow them to live in many different environments. Mammals are a successful group of animals. A group of animals is successful (a) if its members are found in many types of environments, (b) if there are many of them or (c) if they have survived for long periods of evolutionary time.

One trait that allows mammals to live in many different environments is that they have developed special ways of protecting their young from the dangers of the environment. One way that the young are protected is that fertilization is internal in mammals, and the young develop inside the mother's body. The amount of development at birth varies with the type of mammal. A mammal that is not well developed at birth is more dependent on its parents' care than one that is fully developed. Mammals care more for their young than other kinds of animals. The mammary gland is found only in mammals and is used in the care of their young. A mammary gland is a structure in female mammals that secretes milk. The milk provides food for the young.

Another trait that allows mammals to live in many different places is that they are warm-blooded. Unlike reptiles, such as snakes or lizards, which need the sun to stay warm, the mammal's body can maintain a nearly constant temperature, regardless of the temperature of the environment. Like birds, they have a four-chambered heart that helps them do this. The heart keeps the body warm by pumping blood to all parts of the body.

A third trait of mammals that allows them to live in many different environments is that they are the only animals that have hair or fur. Though some mammals have little or no hair, most are covered with hair. Hair insulates mammals in much the same way that feathers insulate birds. It keeps them from getting too hot or too cold.

Another physical trait of mammals is that they can eat many different kinds of food because they have very specialized teeth. This trait also helps them to live in different kinds of environments. There are four types of teeth in mammals: incisors, canines, premolars, and molars. The number and shape of each of these types of teeth are related to the kind of food the mammal eats. Meat-eating mammals, such as wolves and lions, have long, pointed canine teeth, that are used for tearing. Their incisors are chisel-shaped and are used for cutting. Plant-eating mammals, such as horses and cows, have large, flat premolars and molars. These teeth are used for grinding plant materials.

Another important reason why mammals are so successful is that they have a well-developed nervous system that includes a complex brain. In general, the brain of a mammal is larger than that of other vertebrates. Mammals are more intelligent than most other vertebrates. Their large and complex brain supports the complex behaviour which is another characteristic trait of mammals.

Mammals have two types of behaviour: learned and instinctive behaviours. The ability to learn complex behaviours also contributes to the success of mammals. An example of learned behaviour is seen in bears. Alaskan brown bears teach their offspring to hunt salmon. The young bears carefully watch their mother. After practice, the young bears will become skilled hunters.

Instincts are complex, inborn patterns of behaviour that don't have to be learned. For example, soon after a mammal is born, it finds its mother's mammary gland, or breast, and begins to feed on milk. Without being taught, the newborn seeks its mother's breast. This behaviour is inborn. Breast-feeding is one type of instinct that mammals have, but mammals have many other types of instincts. Migration is an instinct of some mammals. The defense of certain territory is another. Hibernation is another instinct of mammals. Hibernation is a type of deep sleep in which an animal has a lowered body temperature. Many mammals hibernate during winter. Food, needed to supply energy, is scarce in winter. A great amount of energy is needed to maintain a normal body temperature during cold weather. Since a hibernating animal maintains a low body temperature, it uses little energy. Thus, mammals who hibernate can live in environments with cold winters.

Question Stems

- What do you think might occur if.....? answer:.....
- What information do we already have about.....?
How does it apply to? answer:.....
- Are there any differences between.....and....? Explain.
-appears to be a problem because.....
What are some possible solutions?
- The author(s) states that "....."
Explain why this statement is true or false.
- Compare.....and/within regards to.....
Explain your answer.

Signal Words

Who.....?

What.....?

Where.....?

When.....?

Why.....?

How.....?

QUESTION STEM EXAMPLES

1. **What do you think might occur if a mammal developed outside of its mother's body? What dangers would there be for the baby's survival?**

answer:.....

2. **What information do we already have about reptiles versus mammals? How do these differences apply to body temperature?**

3. **Are there any differences between learned and instinctive behaviour? Explain.**

4. **The untimely death of a mother appears to be a problem because of her role as nurturer and teacher. What are some possible solutions for the baby's survival if the mother should die prematurely?**

5. **The author(s) states that "Mammals...are the only animals that have hair or fur." Explain why this statement is true or false.**

6. **Compare incisors and canines with premolars and molars In regards to eating. Explain your answer.**

SIGNAL WORD QUESTION EXAMPLES

Memory Questions:

1. Who is included in the list as being a mammal?
2. What do mammals eat?
3. Where do baby mammals develop before being born?
4. When do mammals hibernate?
5. Why do mammals have hair?
6. How do reptiles stay warm?

Critical Thinking Questions:

1. Who is the parent of primary importance for a young mammal? Explain.
2. What other species not listed in the text are classified as mammals? Explain.
3. Where are mammals found and how are they able to live in these environments?
4. When do mammals learn their specialized behaviours?
5. Why does a meat-eating mammal have different teeth as compared to a plant-eating mammal?
6. How do mammals survive through the winter when hibernating?

Peru and Argentina

Many factors influence whether a developing country can look forward to a prosperous future, or whether it will be forever doomed to backwardness and poverty. A comparison between Peru and Argentina illustrates this point.

The lofty, snowcapped peaks of the Andes mountains cover most of Peru. Even the valleys are so high that the air is thin and cold. Where the mountains reach down to the Pacific Ocean, it never rains, and the air is so dry that even wooden tools and cloth made hundreds of years ago are perfectly preserved in the sand.

In contrast, the heartland of Argentina consists of huge, grassy plain, called the Pampas, where cowboys herd cattle on ranches as large as those in Texas. Where rivers flow into the Atlantic, the capitol city sprawls, with stately buildings and broad avenues, almost like in Paris.

Wheat fields cover much of the Pampas, whose soil is so rich that it never needs fertilizing. However, Argentina, like America, has trouble finding buyers for all the wheat and beef it produces. The country has almost no coal, iron or other minerals. Thus it cannot manufacture many things that people need, such as cars, machinery, and clothing, so these items must be imported from other countries.

Peru is very different. There are many small farmers who grow just enough corn to feed their families. Other people work on large plantations where coffee, cotton, and sugar cane is produced for export to other countries. The mountains have rich deposits of copper, silver, and lead, and the government has developed some very profitable mines. The Indians who live in the high mountain valleys raise sheep and llamas. The women weave beautiful sweaters and blankets from the wool of the llamas, just as their ancestors did. These weavings eventually find their way to the fashionable boutiques of Europe and the U.S. where they bring high prices.

Most Peruvians are Indians, descendants of the once proud and mighty Incas. When the Spanish destroyed the Inca empire, the Indians became the poor and oppressed people in their own land. A small group of white plantation owners has ruled the country ever since. They brought with them Christianity and the Spanish language, but the great Indian masses remained outside the Spanish culture. Most of them do not even speak Spanish, and few can read or write.

In Argentina, on the other hand, there are very few Indians today. The streets in the big cities are crowded with busy, energetic people, and in the outdoor cafes lively discussions can be heard. In recent years, the newspapers and magazines have been free to publish everything, for the people were finally able to establish a free, democratic government.

Thus, Peru and Argentina may develop very differently in the future.

APPENDIX D

Strategy Prompts

Question Stems

Instructions:

1. You are to generate **6 questions** using the generic question stems.
 2. You are to use each question stem **only once**.
 2. Be sure to **answer** your question.
 3. If you are pressed for time, create your question and return to answer it later.
 4. You may refer to and create links with the text entitled "Blood", **but your task is to create questions using the "Heart Disease" text.**
 5. You have 20 minutes.
-

- What do you think might occur if.....? answer:.....
- What information do we already have about.....?
How does it apply to? answer:.....
- Are there any differences between..... and? Explain.
-appears to be a problem because.....
What are some possible solutions?
- The author(s) states that "....."
Explain why this statement is true or false.
- Compare.....and/within regards to.....
Explain your answer.

Signal Words

Instructions:

1. You are to generate **6 questions** using the signal words.
 2. You are to use each signal word **only once**.
 3. Be sure to **answer** your question.
 4. If you are **pressed for time**, create your question and return to answer it later.
 5. You may refer to and create links with the text entitled "Blood", but your **task is to create questions using the "Heart Disease" text**.
 6. You have 20 minutes.
-

Who.....?

What.....?

Where.....?

When.....?

Why.....?

How.....?

Questioning Reminder

1. Carefully read the text entitled **"Heart Disease"**.
2. Once you feel that you understand what you have read, create **6 questions**.
(Suggestion: To help you think of good questions, pretend that you are going to ask another student questions about what you have just read.)
3. Be sure to **answer** your question.
4. If you are pressed for time, create your question and return to answer it later.
5. You may refer to and create links with the text entitled **"Blood"**, but **your task is to create questions using the "Heart Disease" text**.
6. You have 20 minutes.

APPENDIX E

Heart Disease Text (High Coherence)

Heart Disease

The heart is the hardest-working organ in the body. We rely on it to supply blood regularly to the body every moment of every day. Any disorder that stops the heart from supplying blood to the body is a threat to life. Heart disease is such a disorder. It is very common. More people are killed every year in the U.S. by heart disease than any other disease.

There are many kinds of heart disease, some of which are present at birth and some of which are acquired later.

1. Congenital heart disease

A congenital heart disease is a defect that a baby is born with. Most babies are born with perfect hearts. But one in every 200 babies is born with a bad heart. For example, hearts have flaps, called valves, that control the blood flow between its chambers. Sometimes a valve develops the wrong shape. It may be too tight, or fail to close properly, resulting in congenital heart disease. Sometimes a gap is left in the wall, or septum, between the two sides of the heart. This congenital heart disease is often called a "septal defect". When a baby's heart is badly shaped, it cannot work efficiently. It cannot pump enough blood through the lungs so that it receives enough oxygen. As a result, the baby becomes breathless. The blood also cannot get rid of carbon dioxide through the lungs. Therefore, the blood becomes purplish, which causes the baby's skin to look blue. Fortunately, it is now possible to save the lives of many "blue babies".

2. Acquired heart disease

Some heart diseases are acquired after the baby is born. Rheumatic fever is an example of an acquired disease that may cause damage to the heart. This disease usually follows a sore throat caused by bacteria known as streptococci. This is often called "strep throat". When strep throat causes rheumatic fever, the tissues of the heart become inflamed. If the heart is badly affected, it fails very soon. Usually, however, it recovers, and the results of the damage are seen only years later. This is because the rheumatic fever leaves scars in the valves of the heart. Therefore, they cannot work properly. This puts a strain on the heart so that eventually it may fail. The effects of the rheumatic fever may take up to twenty or thirty years to appear.

Coronary disease is another example of an acquired heart disease. This disease affects the coronary arteries. These are the blood vessels that extend across the heart and supply it with blood from the lungs. They are very important because they give the heart muscle the oxygen it needs to carry on working. In coronary disease the coronary arteries become blocked, causing

parts of the heart muscle to die because of the lack of oxygen. When this happens, the patient has a heart attack, which can be fatal. The blockage of a coronary artery is usually caused by a clot of blood, called a "thrombus". When a clot forms in a coronary artery, this is called "coronary thrombosis". That is the correct name for a heart attack. In normal arteries, blood does not form clots. But in coronary disease, the walls of the arteries are not normal. They become lumpy, rough, and narrow. The lumps break off and form clots that stop the flow of blood to the heart.

Other examples of acquired heart disease are arrhythmia, angina, and high blood pressure. Arrhythmia, which means "lack of rhythm", is an interruption of the heart's normal beat. Angina is a sharp pain in the chest which is very similar to that caused by a heart attack, or thrombosis. High blood pressure is one of the most common heart diseases. It places a heavy strain on the heart and other organs. Therefore, it is not treated, high blood pressure may lead to heart attacks, kidney failure, or other serious problems. High blood pressure is a disease which has no symptoms. Thus, a person may not be aware of having it unless the blood pressure is measured.

3. Treatment and prevention of heart disease

Since the mid-1960's, medical science has made tremendous progress in the treatment and prevention of heart disease. Both new drugs and new surgical methods have been developed. Among the new drugs for treating heart disease are chemicals called "beta blockers". The beta-blockers lessen the after-effects of heart attacks; they can prevent second attacks; and they can lower the blood pressure of people who have high blood pressure. Other drugs dissolve the lumps which break off the walls of arteries so that they do not stop the flow of blood to the heart.

Surgical techniques for treating heart disease range from repairing or replacing damaged parts, such as valves or arteries, to replacement of the entire heart. If a heart has been so damaged that it can no longer function, it can be replaced by mechanical heart, or, more often, by a heart transplant. In transplant surgery, the healthy heart of someone who has died replaces the diseased heart of the patient. Mechanical devices can be implanted in people's bodies to keep their hearts functioning. The pacemaker is the most common of these devices. It does not heal the diseased heart, but it relieves the symptoms of an irregular heart beat and maintains the steady beat needed for normal living. When a heart cannot pump enough blood through the lungs because of poorly functioning valves, the valves can be replaced with artificial ones of plastic and metal. For patients with coronary disease, "by-pass surgery" is often used to repair clogged or damaged arteries. Doctors use pieces of a patient's own veins, often from the leg, to replace the damaged portions of arteries.

Preventive care is also getting better as scientists learn more and more about the causes of heart disease. They have shown that diet can be an important means of controlling heart disease. For example, a substance called cholesterol is known to cause a build-up of fatty substances in the blood vessels, which can cause blood clots to form in the arteries. Therefore, doctors stress the importance of a diet low-salt diet for patients with high blood pressure.

APPENDIX F

Heart Disease Text (Low Coherence)

Heart Disease

The heart is the hardest-working organ in the body. We rely on it to supply blood regularly to the body every moment of every day. Any disorder that stops the heart from supplying blood to the body is a threat to life. Heart disease is very common. More people are killed every year in the U.S. by heart disease than any other disease.

A congenital heart disease is one that a person is born with. Most babies are born with perfect hearts. In about one in every 200 cases something goes wrong. Sometimes a valve develops the wrong shape. It may be too tight, or fail to close properly. Sometimes a gap is left in the septal wall between the two sides of the heart. This is often called a septal defect. When a baby's heart is badly shaped, it cannot work efficiently. The blood does not receive enough oxygen. The baby becomes breathless. The blood cannot get rid of carbon dioxide through the lungs. It becomes purplish, and the baby's skin looks blue. It is now possible to save the lives of many blue babies.

The disease called rheumatic fever may cause damage to the heart. The disease usually follows a sore throat caused by bacteria known as streptococci. The tissues of the heart become inflamed. If it is badly affected, it fails. Usually it recovers, and the results of the damage are seen only years later. The valves of the heart are left with scars. They cannot work properly. This puts a strain on the heart. Eventually it may fail. The effects of the rheumatic fever may take up to twenty or thirty years to appear.

The blood vessels that extend across the heart and supply it with blood are called coronary arteries. They are very important. They give the heart the oxygen it needs to carry on working. If they become blocked, parts of the heart muscle will die. The patient has a heart attack, which can be fatal. The blockage of a coronary artery is usually caused by a thrombus, or blood clot. Coronary thrombosis happens when a clot forms in a coronary artery. That is the correct name for a heart attack. In normal arteries, blood does not form clots. In coronary disease, the walls of the blood vessels become lumpy, rough, and narrow.

Arrhythmia is an interruption of the heart's normal beat. Angina is a sharp pain in the chest which is very similar to that caused by thrombosis. High blood pressure is very common. If untreated, high blood pressure may lead to heart attacks, kidney failure, or other serious problems. High blood pressure may have no symptoms. A person may not be aware of having it unless the blood pressure is measured.

Among the new drugs for treating heart disease are a family of

compounds called beta blocking drugs, or simply, beta blockers. They lessen the after-effects of heart attacks, can prevent second attacks, and can lower the blood pressure of people who have high blood pressure. Other drugs dissolve the lumps which break off the walls of veins and arteries.

Heart transplants are used more often than mechanical hearts. In transplant surgery, the healthy heart of someone who has died replaces the heart of the patient. Mechanical devices can be implanted in people's bodies to keep their hearts functioning. The most commonly used pacemaker does not heal the diseased heart, but it relieves the symptoms of an irregular heart and keeps a steady beat for normal living. When a heart cannot pump enough blood through the lungs because of poorly functioning valves, the valves can be replaced with artificial ones of plastic and metal. By-pass surgery is used to repair clogged or damaged blood vessels. Doctors use pieces of a patient's own veins, often from the leg, to replace the damaged portions of arteries.

A substance called cholesterol is known to cause a build-up of fatty substances in the blood vessels, which can lead to heart disease, so doctors stress the importance of a diet low in fats. Salt is known to increase the blood pressure, so a low-salt diet is recommended.

APPENDIX G

Blood Text

Blood

Blood is the river of life that flows within us, transporting nearly everything that must be carried in the body from one place to another. Therefore, it has extremely important functions which keep the body alive and healthy. To understand the importance of blood, it is necessary to examine its components and how it functions within the body.

1. Components of blood

Among all of the body's tissues, blood is unique. It is the only tissue in the body which is a fluid. At first glance it may appear that blood is a simple fluid. In truth, it is a very complex mixture of solid and liquid components or parts. Essentially, blood is a complex tissue in which living blood cells, known as the formed elements, are suspended in a nonliving fluid called plasma.

If a sample of blood is spun in a centrifuge (a machine designed to separate liquids through rapid spinning), the heavier formed elements are packed down by force (centrifugal force) and the lighter plasma rises to the top (Figure 1). Most of the reddish mass at the bottom of the tube consists of erythrocytes, the red blood cells that help to transport oxygen in the body. A thick, whitish layer called the buffy coat lies between the plasma and erythrocyte. This layer contains leukocytes (white blood cells), which help protect the body's immune system, and platelets, or cell fragments that help the blood to clot.

2. Physical characteristics and volume

Blood is a sticky fluid with a characteristic salty taste. Depending upon the amount of oxygen it is carrying, the color of blood varies from scarlet (oxygen-rich) to a dark red (oxygen-poor). Blood is heavier than water and about five times thicker. It accounts for 8% of body weight. Finally, blood is slightly alkaline, with a pH between 7.35 and 7.45 (7.00 is normal, anything below that number is acidic and anything above is alkaline).

3. Functions of blood in the body

Blood performs a number of functions which interact and overlap with one another for proper body maintenance. All are concerned in one way or another with protecting the body and distributing substances necessary for good health.

a. Distribution of substances:

- i. Blood delivers both oxygen from the lungs and nutrients from the digestive system to all cells within the body.
- ii. Blood transports waste products from body cells to those organs which eliminate toxins (i.e., transporting blood to the lungs for

elimination of carbon dioxide and to the kidneys for elimination of nitrogenous wastes in urine).

- iii. Blood transports hormones to the body's organs.
- iv. Blood helps to maintain a healthy body temperature through the absorption and distribution of body heat.

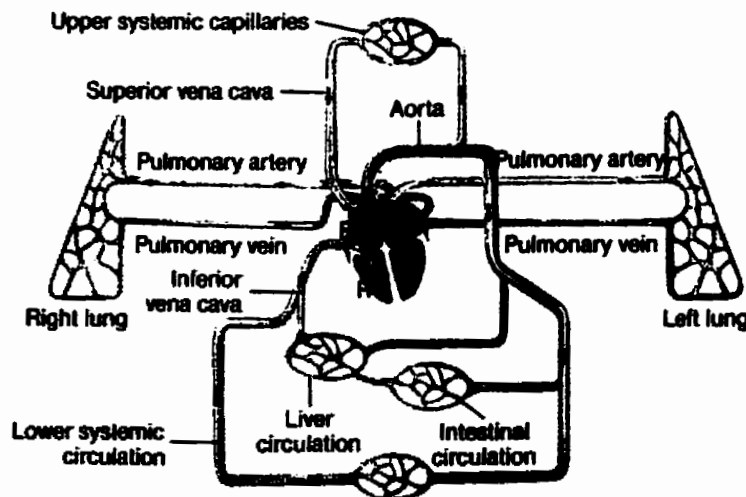
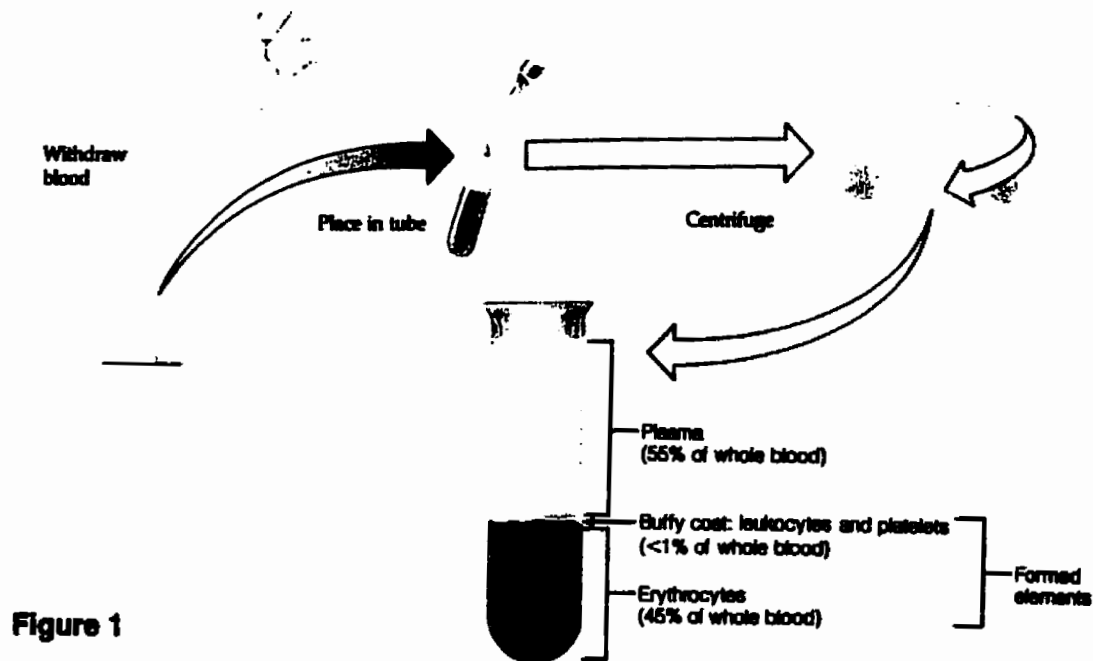
b. Protection of the body:

- i. Blood helps to maintain a normal pH level in all body tissues. Many proteins in the blood and blood solutes help to keep pH levels stable so that body cells can perform normally. If blood is too acidic or too alkaline, body cells fail to function properly.
- ii. Blood works to ensure that there is an adequate supply of itself in the circulatory system. Chemicals found in the blood work together to ensure that blood vessels remain full and that there are no leaks from the blood stream into tissue spaces. This is important as full blood vessels are necessary for proper blood flow throughout the body.
- iii. The components of blood work together to prevent blood loss. When a blood vessel is damaged, platelets and plasma proteins initiate clotting.
- iv. Blood helps to prevent infection within the body. Drifting along in the blood are antibodies, proteins, and white blood cells, all of which help defend the body against foreign invaders such as bacteria, viruses, toxins, and tumor cells.

4. Blood vessels

It is evident that blood must continuously move throughout the body in order to maintain proper health. The way blood moves is through blood vessels. These vessels form an elaborate system of delivering blood that begins and ends at the heart. There are three major types of blood vessels: arteries, capillaries, and veins (Figure 2). As the heart pumps and relaxes, blood is forced out of the heart into the large arteries of the body. From there it moves into successively smaller arteries, finally reaching the smallest branches, or arterioles which feed into the capillary beds of all body organs and tissues. Blood draining from the capillaries is then collected by venules, or small veins that merge to form larger veins which ultimately empty into the great veins that return back to the heart. Altogether, blood vessels carry blood on a journey that stretches for about 96,558 kilometers (60,000 miles) through an adult body.

Blood



The right chamber receives oxygen-depleted blood from the circulation. Blood is pumped out of the RV, routed through the lungs, where the oxygen is replenished, and returned to the left side of the heart. From the LV, blood is propelled into the systemic circulation.

Figure 2

APPENDIX H

Sorting Measure

CATEGORY SORT

Identification # _____

Written below are 18 words which you are being asked to sort into categories.

1. You may put the words into **any** category which you think is appropriate.
2. You may make **as few or as many** categories as you wish.
3. You are **not required** to have the **same number of words** in each of the categories.
4. You may use a word **more than once**.
5. You may **change your mind and reorganize the words** at any time. Simply draw a line through your response indicating that you do not want it included in that category.
6. There is **no correct or incorrect way** to organize the words.
7. At the end of each category list **briefly state why those words go together**.
8. Give each category a name.

MULTIPLE SCLEROSIS	PULMONARY VEIN	ARRHYTHMIA	SEPTAL DEFECT
VENTRICLE	BY-PASS SURGERY	CANCER	PACEMAKER
BLUE BABY	MITRAL VALVE	STREPTOCOCCI	CARBON DIOXIDE
BLOOD CLOT	MALIGNANT	THYROID	RHEUMATIC FEVER
KIDNEY	CORONARY THROMBOSIS		

Place responses here.

Sorting Matrix

Ideal Sorting Matrix

blue baby/septal defect	1
blue baby/carbon dioxide	1
septal defect/carbon dioxide	1
blood clot/coronary thrombosis	1
blood clot/bypass surgery	1
coronary thrombosis/bypass surgery	1
rheumatic fever/streptococci	1
arrhythmia/pacemaker	1
bypass surgery/pacemaker	1
ventricle/mitral valve	.5
ventricle/pulmonary vein	.5
mitral valve/pulmonary vein	.5
thyroid/kidney	.5
cancer/multiple sclerosis	.5
cancer/malignant	.5
multiple sclerosis/malignant	.5
	<hr/> 12.5

IDEAL SORTING MATRIX (WEIGHT MATRIX)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0																	
2	1	0																
3	1	1	0															
4	-.0456	-	-	0														
5	-	-	-	1	0													
6	-	-	-	1	1	0												
7	-	-	-	-	-	-	0											
8	-	-	-	-	-	-	1	0										
9	-	-	-	-	-	-	-	-	0									
10	-	-	-	-	-	1	-	-	1	0								
11	-	-	-	-	-	-	-	-	-	-	0							
12	-	-	-	-	-	-	-	-	-	-	.5	0						
13	-	-	-	-	-	-	-	-	-	-	.5	.5	0					
14	-	-	-	-	-	-	-	-	-	-	-	-	-	0				
15	-	-	-	-	-	-	-	-	-	-	-	-	-	.5	0			
16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0		
17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.5	0	
18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.5	.5	0

KEY

- | | |
|------------------------|------------------------|
| 1. BLUE BABY | 11. VENTRICLE |
| 2. SEPTAL DEFECT | 12. MITRAL VALVE |
| 3. CARBON DIOXIDE | 13. PULMONARY VEIN |
| 4. BLOOD CLOT | 14. THYROID |
| 5. CORONARY THROMBOSIS | 15. KIDNEY |
| 6. BYPASS SURGERY | 16. CANCER |
| 7. RHEUMATIC FEVER | 17. MULTIPLE SCLEROSIS |
| 8. STREPTOCOCCI | 18. MALIGNANT |
| 9. ARRHYTHMIA | |
| 10. PACEMAKER | |

By multiplying two matrices or vectors, that is, calculating the inner products of the two, one derives a statistical computation of the similarity, or harmony, of the two matrices. Thus, the sort score is the sum of the inner product between a participant's sorting matrix (s_{ij}) and a weight matrix (w_{ij}) divided by the sum of the positive weights in the weight matrix ($\sum w_{ij} / 2$): $((\sum (s_{ij} \times w_{ij})) / (\sum w_{ij} / 2))$. Specifically, a participant's sorting matrix consists of 1s and 0s, whereby a 1 indicates a pair of items sorted together and a 0 indicates a pair not sorted together. The weight matrix, indicative of an ideal sorting, was constructed by assigning a weight ($w_{ij} = 1$) to cells of the matrix representing pairings between items closely related in the text. A weight ($w_{ij} = .05$) was assigned to pairings between nontext items which were correctly categorized. This yields a total of 25 points for the positive values in the matrix (i.e., 12.5×2). The remaining cells of the weight matrix were assigned negative values in such a way that the sum of the matrix became 0. Specifically, the negative weight was set equal to $-((\sum w_{ij} / 2) / (m - k))$, where $\sum w_{ij} / 2$ is the sum of the positive values in the matrix, m is the total number of nondiagonal cells in the matrix $n^2 - n$, and k is the number of cells containing positive values. Diagonal cells of the weight matrix were assigned a weight of 0."

$$w/2 = 25/2$$

$$12.5/274 = -.0456204$$

$$n^2 = 324$$

$$n = 18$$

$$k = 32$$

PARTICIPANT'S SORTING MATRIX

SAMPLE A

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0																	
2	0	0																
3	1	0	0															
4	0	1	0	0														
5	0	1	0	1	0													
6	0	1	0	1	1	0												
7	1	0	1	0	0	0	0											
8	1	0	1	0	0	0	1	0										
9	0	1	0	1	1	1	0	0	0									
10	0	1	0	1	1	1	0	1	1	0								
11	0	0	0	0	0	0	0	0	0	0	0							
12	0	0	0	0	0	0	0	0	0	0	0	.5	0					
13	0	0	0	0	0	0	0	0	0	0	0	.5	.5	0				
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.5	0		
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
17	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.5	0	0

KEY

1. BLUE BABY
2. SEPTAL DEFECT
3. CARBON DIOXIDE
4. BLOOD CLOT
5. CORONARY THROMBOSIS
6. BYPASS SURGERY
7. RHEUMATIC FEVER
8. STREPTOCOCCI
9. ARRHYTHMIA
10. PACEMAKER

11. VENTRICLE
12. MITRAL VALVE
13. PULMONARY VEIN
14. THYROID
15. KIDNEY
16. CANCER
17. MULTIPLE SCLEROSIS
18. MALIGNANT

$$9.5 + 18(-.0456204) = .694$$

12.5

APPENDIX I

Calibration of Confidence Measure

CONFIDENCE MEASURE

Identification # _____

Now that you have read the text entitled "Heart Disease", please answer the following questions as honestly and as carefully as possible.

After you have read each question, please circle a single number on the accompanying scale to report how confident you are that you could correctly answer the question.

1. Could you discuss the condition known as "strep throat" and how it should be treated?

1	2	3	4	5	6	7
not at all confident						extremely confident

2. Could you explain the purpose of "by-pass" surgery?

1	2	3	4	5	6	7
not at all confident						extremely confident

3. Could you explain why blood must flow properly between the two sides of the heart for the body to stay healthy?

1	2	3	4	5	6	7
not at all confident						extremely confident

4. Could you explain what happens when coronary arteries become blocked?

1	2	3	4	5	6	7
not at all confident						extremely confident

5. Could you explain the impact of an unhealthy lifestyle on a person's heart?

1	2	3	4	5	6	7
not at all confident						extremely confident

6. Could you explain how a pacemaker helps a diseased heart to function better?

1	2	3	4	5	6	7
not at all confident						extremely confident

7. Could you explain why
arrhythmia is a threat to life?

1	2	3	4	5	6	7
not at all confident						extremely confident

8. Could you recommend an alternate
treatment to the use of beta blockers?

1	2	3	4	5	6	7
not at all confident						extremely confident

9. Could you explain how regular
exercise is good for the heart and
for blood circulation?

1	2	3	4	5	6	7
not at all confident						extremely confident

10. Could you explain what causes
the condition "blue baby" ?

1	2	3	4	5	6	7
not at all confident						extremely confident

11. Could you discuss the recommended
dietary changes known to control heart
disease?

1	2	3	4	5	6	7
not at all confident						extremely confident

12. Could you explain the consequences
of high blood pressure?

1	2	3	4	5	6	7
not at all confident						extremely confident

APPENDIX J
Post-Test Measure

Post-Test

Identification # _____

Answer the following questions as carefully and as fully as possible. You may use point form, but be sure that your responses are clear and concise.

1. What is a pacemaker? Explain how it helps a patient with heart disease to survive.
2. What is the purpose of "by-pass" surgery?
3. A 62 year old man has had two heart attacks in five years. He was taking beta blockers after his first heart attack. Why should the doctor consider a new treatment and what should it be? Be specific.
4. What happens when the coronary arteries become blocked?
5. What causes the condition known as "blue baby"?
6. If left untreated, what are the results of high blood pressure?
7. You are the parent of a 6 year old child. For the past two days, your child has complained of a very sore throat. You suspect that it may be strep throat since several of the children in your child's classroom have been diagnosed with the condition. What is strep throat, how should it be treated, and what are some possible short-term and long-term consequences if it is not?
8. Explain why arrhythmia is a threat to life.
9. Explain how regular exercise is good for the heart and for blood circulation.
10. When valves in the heart are scarred from rheumatic fever they cannot function properly. Explain why blood must flow properly between both sides (i.e., the chambers) of the heart in order for the body to stay healthy.
11. What are the recommended dietary changes known to control heart disease? Explain why they are beneficial.
12. You are a doctor. Your patient is a 56 year old man. He has an extremely stressful job, he smokes, he eats a high fat diet, and he has history of heart disease in his family. He does not want to change his lifestyle. Explain in as much detail as possible what each one of these factors is doing to his heart and overall health.

Post-test Answer Key

Scoring Procedures:

7 = if all of the underlined information is present in the answer including a clear understanding of the subject (i.e., elaborative or additional details, a detailed or knowledgeable description of the information, etc.).

6 = if all of the underlined information is present in the answer without including a clear understanding of the subject (i.e., elaborative or additional details, a detailed or knowledgeable description of the information, etc.).

5 = if most (3/4) of the underlined information is present in the answer including a clear understanding of the subject (i.e., elaborative or additional details, a detailed or knowledgeable description of the information, etc.).

4 = if most (3/4) of the underlined information is present in the answer without including a clear understanding of the subject (i.e., elaborative or additional details, a detailed or knowledgeable description of the information, etc.).

3 = if some (1 or 2 points, depending on the length of the answer) of the underlined information is present in the answer including a clear understanding of the subject (i.e., elaborative or additional details, a detailed or knowledgeable description of the information, etc.).

2 = if some (1 or 2 points, depending on the length of the answer) of the underlined information is present in the answer without including a clear understanding of the subject (i.e., elaborative or additional details, a detailed or knowledgeable description of the information, etc.).

1 = if an idea or ideas relating to (the gist) the underlined information are present in the answer

0 = if none of the underlined information is present in the answer

Instructions to Students: "Answer the following questions as carefully and as fully as possible. You may use point form, but be sure that your responses are clear and concise".

1. What is a pacemaker? Explain how it helps a patient with heart disease to survive. (Bridging Inference)

A pacemaker is a mechanical device implanted in the body to keep the the heart functioning. While it cannot heal a diseased heart, it can relieve the symptoms of an irregular heart beat and maintain the steady beat needed for normal living. After all, the heart must beat normally and in rhythm so as to pump blood/oxygen through the lungs and back to the heart. If it does not, they will have a heart attack.

2. What is the purpose of "by-pass" surgery? (Textbase)

It is often used to repair clogged or damaged arteries.

3. A 62 year old man has had two heart attacks in five years. He was taking beta blockers after his first heart attack. Why should the doctor consider a new treatment and what should it be? Be specific. (Problem-Solving)

(Because beta-blockers were supposed to prevent a second attack, his problems are obviously much greater than those which beta-blockers can help with.) His doctors could treat his problems with a different drug or they may have to consider some type of surgical technique such as repairing or replacing parts or implanting a mechanical device such as a pacemaker.

4. What happens when the coronary arteries become blocked? (Textbase)

A blockage causes parts of the heart muscle to die because of a lack of oxygen.

5. What causes the condition known as "blue baby"? (Bridging Inference)

When a badly shaped heart cannot pump enough blood through the lungs to receive oxygen, blood becomes full of carbon dioxide. Therefore, the blood becomes purplish which causes a baby's skin to look blue.

6. If left untreated, what are the results of high blood pressure? (Textbase)

It may lead to heart attacks, kidney failure, or other serious problems.

7. You are the parent of a 6 year old child. For the past two days, your child has complained of a very sore throat. You suspect that it may be strep throat since several of the children in your child's classroom have been diagnosed with the condition. What is strep throat, how should it be treated, and what are some possible short-term and long-term consequences if it is not? (Problem Solving)

(Strep throat is a sore throat caused by the bacteria streptococci so it is necessary to use antibiotics to impede its development.) If it is not treated, rheumatic fever may result meaning that the tissues of the heart will become inflamed, leaving scars in the valves. This may lead to heart damage which could take 20 to 30 years to manifest. At that time, by-pass surgery will be required to replace those valves.

8. Explain why arrhythmia is a threat to life. (Elaborative Inference)

Arrhythmia is an interruption of the heart's normal beat. When a heart is not beating properly blood is not being pumped to and from the lungs, where it is receiving oxygen (as well as to the rest of the body where it is required for proper organ functioning).

9. Explain how regular exercise is good for the heart and for blood circulation. (Elaborative Inference)

Regular exercise helps the heart muscle to remain strong, as it is a muscle. When the heart is strong, it is better able to pump blood throughout the body (improving one's circulation system). Also, freer flowing blood also means that the body's organs are receiving oxygen, (hormones, and the elimination of toxins).

10. When valves in the heart are scarred from rheumatic fever they cannot function properly. Explain why blood must flow properly between both sides (i.e., the chambers) of the heart in order for the body to stay healthy. (Elaborative Inference)

Because blood enters the heart from the body full of carbon dioxide-laden (deoxygenated) blood and is pumped into the lungs where it receives oxygen. From the lungs blood returns to the heart and is (then pumped back into the body for circulation where it is needed to distribute hormones, eliminate toxins, maintain body temperature, etc).

11. What are the recommended dietary changes known to control heart disease? Explain why they are beneficial. (Bridging Inference)

Doctors stress the importance of a diet low in cholesterol because cholesterol causes a build-up of fatty substances in the blood vessels which causes blood clots in the arteries. Clots prevent oxygen from reaching the heart muscle. Also, doctors recommend a low-salt diet for patients with high blood pressure as salt causes an increase in blood pressure. High blood pressure may lead to a heart attack, kidney failure, or other serious problems.

12. You are a doctor. Your patient is a 56 year old man. He has an extremely stressful job, he smokes, he eats a high fat diet, and he has history of heart disease in his family. He does not want to change his lifestyle. Explain in as much detail as possible what each one of these factors is doing to his heart and overall health. (Problem Solving)

(He is not a young man anymore so his heart is not as resilient as it once was. He has a lot of additional stress which can cause high blood pressure. He smokes so there is less oxygen in his blood stream. He probably has high cholesterol due to a high fat diet. And finally, he is genetically predisposed to heart problems.) Therefore, this man's heart muscle is going to worsen due to all of these factors which means he will have less circulatory capability which will damage his vital organs and body cells. He must given up smoking and his high fat diet and start exercising and going to the doctor regularly for check-ups.

APPENDIX K

Summary Recall Measure

Summary recall

Identification # _____

In the space below, you are being asked to **summarize** what you read in the text entitled **"Heart Disease"**. This means that you are being asked to state the **important information** in the text in a **briefer form**. Please use a paragraph format.

Propositional Analysis

A: AGENT C: TIME AND PLACE G: GOAL
O: OBJECT MOD: MODIFIER

P1 The heart is the hardest-working organ in the body.
IS [A:HEART, O: ORGAN (MOD: HARDEST WORKING), C: BODY]

P2 We rely on it to supply blood regularly to the body every moment of every day.
SUPPLY [A: HEART, O: BLOOD, C: MOMENT (MOD: EVERY),
DAY (MOD: EVERY), G: REGULAR]

P3 Any disorder that stops the heart from supplying blood to the body is a threat to life.
STOPS [A: DISORDER, O: BLOOD SUPPLY, C: BODY, G: LIFE
(MOD: THREAT TO)]

***P4 Heart disease is such a disorder.**
IS [A: HEART DISEASE, O: DISORDER (MOD: SUCH A),
G: LIFE (THREAT TO)]

P5 It is very common.
IS [A: HEART DISEASE, C: COMMON (MOD: VERY)]

P6 More people are killed every year in the U.S. by heart disease than by any other disease.
KILL [A: HEART DISEASE, O: PEOPLE (MOD: MORE), C: YEAR
(MOD: EVERY), U.S., G: DISEASE (MOD: THAN
ANY OTHER)]

****P7 There are many kinds of heart disease, some of which are present at birth and some of which are acquired later.**
ARE [A: HEART DISEASE (MOD:MANY-KIND), A: HEART
DISEASE (MOD: SOME), C: BIRTH (MOD:
PRESENT-AT), C: LATER (MOD: ACQUIRE)]

P8 A congenital heart disease is a {defect} that a baby is born with.
DEFECT [A: HEART DISEASE (MOD: CONGENITAL), O: BABY,
C: BORN (MOD: WITH)]

P9 Most babies are born with perfect hearts.

ARE [A: HEART (MOD: PERFECT), O: BABY (MOD: MOST), C:
BORN]

P10 {But} one in every 200 babies {is born with a bad heart}.

IS [A: HEART (MOD: BAD), O: BABY (MOD: 1/200), C: BORN]

*P11 For example, hearts have flaps, called valves, that control the blood flow between its chambers.

CONTROL [A: HEART, O: VALVE (MOD: FLAP), G:
FLOW (MOD: BLOOD), C: CHAMBER (MOD:
BETWEEN)]

P12 Sometimes a valve develops the wrong shape.

DEVELOP [O: VALVE, G: SHAPE (MOD: WRONG), C:
SOMETIMES]

P13 It may be too tight, or fail to close properly,{ resulting in congenital heart disease}.

RESULT [O: VALVE (MOD: TOO-TIGHT)(MOD: FAIL-TO-CLOSE),
G: HEART DISEASE (MOD: CONGENITAL)]

P14 Sometimes a gap is left in the wall, {or septum}, between the two sides of the heart.

GAP [O: SEPTUM (MOD: HEART-WALL), C: SIDE (MOD:
BETWEEN-TWO), C: SOMETIMES]

P15 {This congenital heart disease} is often called a "septal defect".

CALL [A: HEART DISEASE (MOD: CONGENITAL), G:SEPTAL
DEFECT,C: OFTEN]

P16 When a baby's heart is badly shaped, it cannot work efficiently.

CANNOT-WORK [O: HEART (MOD: BAD-SHAPE)]
(MOD:EFFICIENT)

P17 It cannot pump enough blood through the lungs so that it receives enough oxygen.

CANNOT-PUMP [A: HEART, O: BLOOD (MOD: ENOUGH), G:
LUNGS]

P18 RECEIVE [A: HEART, O: BLOOD, G: OXYGEN (MOD:
ENOUGH)]

- P19 {As a result}, the baby becomes breathless.
BECOME [A: BLOOD, O: BABY, G: BREATHLESS]
- P20 The blood also cannot get rid of carbon dioxide through the lungs.
CANNOT-GET-RID [A: BLOOD, O: CARBON DIOXIDE, G: LUNGS
(MOD: THROUGH)]
- P21 {Therefore, the blood} becomes purplish, which causes the baby's skin to look blue.
CAUSE [A: BLOOD (MOD: PURPLE), O: CARBON DIOXIDE]
[O: BABY, C:SKIN (MOD: BLUE)]
- P22 {Fortunately}, it is now possible to save the lives of many "blue babies".
POSSIBLE [O: LIFE (MOD: BLUE-BABY) , G: SAVE, C: NOW]
- *P23 Some heart diseases are acquired after the baby is born.**
ACQUIRE [A: HEART DISEASE (MOD: SOME), O: BABY,
G: BORN, C: AFTER]
- P24 Rheumatic fever {is an example of an acquired disease that} may cause damage to the heart.
MAY-CAUSE [A: RHEUMATIC FEVER, O: DISEASE (MOD:
ACQUIRE), C: HEART (MOD: DAMAGE)]
- P25 This disease usually follows a sore throat caused by bacteria known as streptococci.
CAUSE [A: STREPTOCOCCI (MOD: BACTERIA), C: THROAT
(MOD: SORE), G: RHEUMATIC FEVER]
- *P26 This is often called "strep throat".**
CALL [A: STREPTOCOCCI, O: STREP THROAT, C:
OFTEN]
- P27 {When strep throat causes rheumatic fever}, the tissues of the heart become inflamed.
CAUSE [A: STREP THROAT, O: RHEUMATIC FEVER]
BECOME [A: RHEUMATIC FEVER, O: HEART (MOD:
INFLAME-TISSUE)]
- P28 If the heart is badly affected, it fails {very soon}.
AFFECT (MOD: BAD) [O: HEART, G: FAIL, C: SOON (MOD: VERY)]

P29 Usually, {however}, it recovers, and the results of the damage are seen only years later.

RESULT [A: RHEUMATIC FEVER, O: HEART (MOD: DAMAGE), C: YEARS-LATER, G: RECOVER (MOD: USUALLY)]

P30 {This is because the rheumatic fever} leaves scars in the valves of the heart.

LEAVE [A:RHEUMATIC FEVER, O: HEART (MOD: VALVE), G: SCAR]

P31 {Therefore}, they cannot work properly.

CANNOT-WORK (MOD: PROPER) [A: RHEUMATIC FEVER, O: VALVE]

P32 This puts a strain on the heart {so that} eventually it may fail.

STRAIN [A: VALVE (MOD: POOR), O: HEART, G: FAIL, C: EVENTUALLY]

P33 The effects of the rheumatic fever may take up to twenty or thirty years to appear.

APPEAR [A: RHEUMATIC FEVER, O: EFFECT, C: YEAR (MOD: 20 - 30)]

****P34 Coronary disease is another example of an acquired heart disease.**

IS [A:CORONARY DISEASE, O: HEART DISEASE (MOD: ACQUIRE), G: EXAMPLE (MOD: ANOTHER)]

***P35 This disease affects the coronary arteries.**

AFFECT [A: CORONARY DISEASE, O: CORONARY ARTERY]

P36 These are the blood vessels that extend across the heart and supply it with blood {from the lungs}.

EXTEND [A: BLOOD VESSELS, O: HEART, G: BLOOD SUPPLY, C: LUNGS (MOD: FROM)]

P37 They are very important because they give the heart muscle the oxygen it needs to carry on working.

GIVE [A: BLOOD VESSELS, O: OXYGEN, C: HEART MUSCLE G: WORK (MOD: CARRY-ON)]

P38 {In coronary disease the coronary arteries} become blocked, causing parts of the heart muscle to die {because of the lack of oxygen}.

{BLOCK [A: CORONARY DISEASE, O: CORONARY ARTERIES] }
CAUSE [A: CORONARY DISEASE, O: HEART MUSCLE
(MOD:PART), G: DIE]
{BECAUSE [A: OXYGEN (MOD: LACK-OF), O:
HEART MUSCLE, G:DIE]}

P39 {When this happens}, the patient has a heart attack, which can be fatal.

HAPPEN [A: OXYGEN (MOD:LACK OF), O: HEART ATTACK
(MOD: PATIENT), G: FATAL]

P40 The blockage of a coronary artery is usually caused by a clot of blood, called a "thrombus".

CAUSE [A: BLOOD CLOT (MOD: THROMBUS), O: CORONARY
ARTERY (MOD: BLOCKAGE), C: USUALLY]

P41 When a clot forms in a coronary artery, this is called "coronary thrombosis".

FORM [A: CLOT, O: CORONARY ARTERY, C: CORONARY
THROMBOSIS (MOD: IN A)]

P42 That is the correct name for a heart attack.

IS [A: CORONARY THROMBOSIS, O: HEART ATTACK, G: NAME
(MOD: CORRECT)]

P43 In normal arteries, blood does not form clots.

DOES-NOT-FORM [A: BLOOD, O: CLOT, C: NORMAL ARTERIES
(MOD: IN)]

P44 {But }in coronary disease, the walls of the arteries {are not normal}.

ARE-NOT-NORMAL [A: CORONARY DISEASE, O:ARTERY, C:
WALL]

P45 {They} become lumpy, rough, and narrow.

BECOME [O: ARTERY, C: WALL (MOD: LUMPY-ROUGH-
NARROW)]

*P46 The lumps break off and form clots that stop the flow of blood to the heart.

BREAK OFF-FORM [A: LUMPS, O: CLOT, G: STOP]
FLOW [A: BLOOD, O: HEART]

****P47 Other examples of acquired heart disease are arrhythmia, angina, and high blood pressure.**

EXAMPLE [A: HEART DISEASE (MOD: ACQUIRE), O: ARRHYTHMIA-ANGINA-HIGH BLOOD PRESSURE]

P48 Arrhythmia, {which means "lack of rhythm"}, is an interruption of the heart's normal beat.

INTERRUPT [A: ARRHYTHMIA (MOD: LACK OF RHYTHM), O: HEART, G: BEAT (MOD: NORMAL)]

P49 Angina is a sharp pain in the chest which is very similar to that caused by {a heart attack, or} thrombosis.

PAIN [A: ANGINA, O: CHEST, C: IN-THE] (MOD: SHARP)

P50 **SIMILAR-CAUSE [A: ANGINA, O: HEART ATTACK (MOD: THROMBOSIS)]**

P51 High blood pressure is {one of the most} common {heart diseases}.

IS [A: HIGH BLOOD PRESSURE, O: HEART DISEASE (MOD: MOST-COMMON)]

***P52 It places a heavy strain on the heart and other organs.**

STRAIN [A: HIGH BLOOD PRESSURE, O: HEART-ORGANS] (MOD: HEAVY)

P53 {Therefore, if it is} not treated, high blood pressure may lead to heart attacks, kidney failure, or other serious problems.

NOT-TREAT [A: HIGH BLOOD PRESSURE, O: HEART ATTACK-KIDNEY FAILURE-SERIOUS PROBLEM, G: LEAD-TO (MOD: MAY)]

P54 High blood pressure {is a disease which has} no symptoms.

IS [A: HIGH BLOOD PRESSURE, O: DISEASE, G: SYMPTOM (MOD: NO)]

P55 {Thus}, a person may not be aware of having it unless the blood pressure is measured.

NOT-AWARE [A: PERSON, O: HIGH BLOOD PRESSURE, G: MEASURE]

*****P56 Since the mid-1960's, medical science has made tremendous progress in the treatment and prevention of heart disease. Both new drugs and new surgical methods have been developed.***

**PROGRESS [A: MEDICAL SCIENCE, O: HEART
DISEASE, G:TREAT-PREVENT, C: MID-1960'S
(MOD: SINCE)] (MOD: TREMENDOUS)
DEVELOP [A: DRUGS-SURGICAL METHODS
(MOD: NEW)]**

P57 Among the new drugs for treating heart disease {are chemicals} called "beta blockers".

**TREAT [A: DRUGS (MOD: NEW), O: CHEMICAL (MOD: BETA
BLOCKERS), G: HEART DISEASE]**

P58 {The beta-blockers} lessen the after-effects of heart attacks; {they} can prevent second attacks; and {they} can lower the blood pressure of people who have high blood pressure.

**LESSEN [A: BETA-BLOCKERS, O: HEART ATTACK (MOD:
AFTER-EFFECT)]**

P59 **PREVENT [A: BETA-BLOCKERS, O: SECOND ATTACK]**

P60 **LOWER [A: BETA-BLOCKERS, O: PEOPLE, G: HIGH
BLOOD PRESSURE]**

P61 Other drugs dissolve the lumps which break off the walls of arteries {so that they do not stop the flow of blood to the heart.}

**DISSOLVE [A: DRUG (MOD: OTHER), O: LUMP (MOD: BREAK-
OFF), C: ARTERY-WALL]
{DO-NOT-STOP [A: BLOOD, O: HEART, G: FLOW, C:
TO-THE]}**

*****P62 Surgical techniques for treating heart disease range from repairing or replacing damaged parts, such as valves or arteries, to replacement of the entire heart.***

**TREAT [A: SURGICAL TECHNIQUE, O: HEART
DISEASE, G: REPAIR-REPLACE, C: VALVE-
ARTERY-HEART]**

P63 {If a heart has been so damaged that it can no longer function}, it can be replaced by a mechanical heart, or, more often, by a heart transplant.

{NOT-FUNCTION [A: DAMAGE, O: HEART]}

REPLACE [A: HEART (MOD: DAMAGE), O: HEART (MOD: MECHANIC-TRANSPLANT), C: MORE-OFTEN]

P64 In transplant surgery, the healthy heart of someone who has died replaces the {diseased} heart of the patient.

REPLACE [A: TRANSPLANT SURGERY, O: HEART (MOD: HEALTHY-DISEASED), C: IN-PATIENT]

P65 Mechanical devices can be implanted in people's bodies to keep their hearts functioning.

IMPLANT [A: DEVICE (MOD: MECHANICAL), O: HEART, G: FUNCTION, C: BODY (MOD: PEOPLE)]

P66 The pacemaker is the most common of these devices.

IS [A: DEVICE (MOD: MOST-COMMON), O: PACEMAKER]

P67 It does not heal the diseased heart, but it relieves the symptoms of an irregular heart beat and {maintains} the steady beat needed for normal living.

DOES-NOT-HEAL [A: PACEMAKER, O: HEART (MOD: DISEASED)]

P68 RELIEVE [A: SYMPTOMS, O: HEART BEAT (MOD: IRREGULAR)]

P69 MAINTAIN [A: BEAT (MOD: STEADY), G: LIFE (MOD: NORMAL)]

P70 When a heart cannot pump enough blood through the lungs because of poorly functioning valves, the valves can be replaced with artificial ones of plastic and metal.

CANNOT-PUMP [A: HEART, O: BLOOD, C: LUNGS (MOD: THROUGH)]

BECAUSE

P71 REPLACE [A: VALVE (MOD: POOR-FUNCTION), O: VALVE (MOD: ARTIFICIAL-PLASTIC-METAL)]

P72 {For patients with coronary disease}, "by-pass surgery" is often used to repair clogged or damaged arteries.

REPAIR [A: BY-PASS-SURGERY, O: CORONARY DISEASE
PATIENT, C: ARTERY (MOD: CLOGGED-DAMAGED)]

P73 Doctors use pieces of a patient's own veins, often from the leg, to replace the damaged portions of arteries.

REPLACE [A: DOCTOR, O: VEIN (MOD: PATIENT-LEG), C:
ARTERY (MOD: DAMAGED)]

****P74 Preventive care is also getting better as scientists learn more and more about the causes of heart disease.**

**LEARN [A: SCIENTIST, O: PREVENTIVE CARE, G:
HEART DISEASE (MOD: CAUSE)]
(MOD: MORE)**

***P75 They have shown that diet can be an important means of controlling heart disease.**

**CONTROL [A: DIET (MOD: IMPORTANT-MEANS), O:
HEART DISEASE]**

P76 {For example}, a substance called cholesterol is known to cause a build-up of fatty substances in the blood vessels, {which can cause blood clots to form in the arteries}.

CAUSE [A: CHOLESTEROL, O: SUBSTANCE (MOD: FAT), C: VESSEL
(BLOOD)]
[A: SUBSTANCE (MOD: FAT), O: CLOT (MOD: BLOOD),
C: ARTERY]

P77 Therefore, doctors stress the importance of a diet low in {cholesterol}.

STRESS [A: DOCTOR, O: DIET, G: CHOLESTEROL (MOD: LOW)]

P78 {Similarly}, salt is known to increase blood pressure, {so doctors} recommend a low-salt diet {for patients with high blood pressure}.

INCREASE [A: SALT (MOD: KNOWN), O: BLOOD PRESSURE, G:
LOW- SALT DIET (MOD: RECOMMEND)]

* These are propositions which are found only in the good microstructure/good macrostructure version of the Heart Disease text.

****These are macropropositions which are found verbatim in the good microstructure/good macrostructure version of the Heart Disease text.**

{ } Information found within these brackets can only be found in the good microstructure/good macrostructure version of the Heart Disease text.

Generalizations, Elaborations, and Reorderings

The summaries are to be scored for the number of text propositions they contain. A liberal, gist scoring criterion should be employed. Nontext statements are to be propositionalized and assigned to inference categories.

Inferences form a continuum in terms of their closeness to the actual text, as opposed to being extrapolations from the reader's own knowledge. However, several categories of inferences can be defined according to the role they play in the comprehension process.

1. **Generalizations** are reductive inferences constructed from more detailed statements in the text. They can be traced to the actual propositions they subsume, except for global generalizations, which are inferences about the overall meaning of the text. Nonetheless, both are generalizations and will be assigned as such. Generalizations reduce the number of text propositions by at least one, although often by many more (E. Kintsch, 1990). For example, several concrete statements about preventative care could be stated as "Humans need to be conscious of their food intake".

2. **Elaborations** are inferences that are not directly implied by the text. Instead, they originate from the subject's own knowledge about the content of the text or related information (E. Kintsch, 1990). For example, "Blood transports waste products from body cells to those organs which eliminate toxins".

This final inference category is to be scored independently of the elaborations and generalizations

3. **Reorderings** are inferences that rearrange text content in an order that is different from the original text. Only between-paragraph (i.e., macrolevel) reorderings are to be considered. There are not scored at the propositional level; instead, an idea in the form of a sentence or paragraph is counted as a reordering if it requires backtracking to an earlier paragraph to find its counterpart in the original text (E. Kintsch, 1990).

Macrostructure levels analysis - Heart Disease

Level 1: Topic and Conclusion

- The subject must mention what the text is about in a broad or general manner. - Example, "The text was about heart disease" "The text was about congenital and acquired heart disease"
- Score each response as 1 point
 - 1 - label (it's about heart disease)
 - 2 - heart disease is both congenital and acquired

Level 2 : Subtopics

- The subject must explicitly mention these 3 inferred or explicit macropropositions in order to be judged as correct. Each statement must be written so as to be deemed a macroproposition, not a level 3 or level 4 response.
- Example, "Congenital heart disease is something a person is born with" "Heart disease can also be acquired" "Much has been discovered about the treatment and prevention of heart disease"
- Score each response as 1 point.
 - 1 - congenital heart disease
 - 2 - acquired heart disease
 - 3 - treatment and prevention of heart disease

Level 3: Subheadings

- The subject must mention one of these text-based subheadings prior to providing detailed information.
- Example, "Sometimes a baby is born with misshaped valves" "Rheumatic fever is an acquired disease which can lead to heart problems" "Heart disease can be treated with drugs"
- Score each response as 1 point.

congenital

- 1 - misshaped valves
- 2 - septal defect
- 3 - blue babies

acquired

- 4 - rheumatic fever
- 5 - coronary disease/thrombosis
- 6 - arrhythmia
- 7 - angina
- 8 - high blood pressure

- treatment
- 9 - drugs (i.e. beta-blockers, others that dissolve lumps)
 - 10 - surgical techniques (i.e. by-pass surgery, transplant, mechanical, pacemaker)
 - 11 - preventative care/diet (i.e. cholesterol, salt)

Level 4: Details

- The subject must include most of the detail as outlined in the proceeding responses to warrant a correct response.
- Example, "Most babies are born with a perfect heart" "Approximately 1 out of 200 babies is born with a bad heart"
- Score each response as 1 point.

- congenital
- 1 - most are born with a perfect heart
 - 2 - 1/200 born with a bad heart
 - 3 - valve is too tight or fails to close properly
 - 4 - gap left in wall/septum between two sides of heart/septal defect
 - 5 - misshaped heart/fails to pump blood through lungs for oxygen
 - 6 - blood is purple when too much carbon dioxide
 - 7 - now possible to save lives of blue babies

- acquired
- 8 - streptococci/strep throat causes rheumatic fever
 - 9 - streptococci inflames heart tissues
 - 10 - heart failure right away or years later (20 to 30 years)
 - 11 - rheumatic fever scars heart valves
 - 12 - coronary arteries are blood vessels that supply heart with necessary oxygen from the lungs
 - 13 - coronary disease is blocked coronary arteries/causes heart attack because of lack of oxygen
 - 14 - blood clot in coronary artery is a thrombus
 - 15 - coronary thrombosis is the correct name for a heart attack
 - 16 - diseased arteries have lumpy, rough, and narrow walls
 - 17 - arrhythmia - interruption in normal heart beat/lack of rhythm
 - 18 - angina - sharp pain in the chest
 - 19 - high blood pressure very common
 - 20 - puts heavy strain on vital organs (i.e., heart attack, kidney, etc)
 - 21 - has no symptoms
 - 22 - patient may be unaware unless it is measured

- treatment**
- 23 - new drugs for treating heart disease are chemicals/compounds called beta blockers**
 - 24 - beta blockers - lessen after-effects of heart attacks**
 - 25 - prevent second attacks**
 - 26 - lower high blood pressure**
 - 27 - other drugs dissolve lumps in arteries to improve blood flow**
 - 28 - transplant surgery replace patient's diseased heart with a healthy heart**
 - 29 - mechanical devices can be implanted to keep heart functioning**
 - 30 - pacemaker is the most common**
 - 31 - relieves symptoms of irregular heart beat and maintains steady beat**
 - 32 - can replace valves with artificial ones (i.e., plastic or metal)**
 - 33 - by-pass surgery to repair clogged or damaged arteries**
 - 34 - use veins from patient (i.e., leg) to replace damaged portions**
 - 35 - diet controls heart disease**
 - 36 - cholesterol is a fatty substance in blood vessels that cause blood clots in the arteries**
 - 37 - low cholesterol diet is stressed**
 - 38 - salt increases blood pressure**
 - 39 - doctors recommend a low salt diet when a patient has high blood pressure**

APPENDIX L
University Advisory Committee on Ethics
in Behavioural Science Research Consent Letter



**UNIVERSITY ADVISORY COMMITTEE
ON ETHICS IN BEHAVIOURAL SCIENCE RESEARCH**

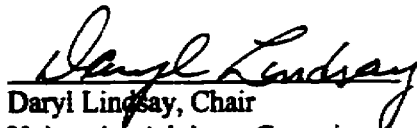
NAME: B.R. Randhawa (T. Gunn)
Department of Educational Psychology
And Special Education

BSC #: 1998-27

DATE: February 1, 1999

The University Advisory Committee on Ethics in Behavioural Science Research has reviewed the Application for Ethics Approval for your study "The Effects of Cross-Modality Processing and Questions Construction on Expository Text Comprehension" (98-27).

1. Your study has been APPROVED.
2. Any significant changes to your proposed study should be reported to the Chair for Committee consideration in advance of its implementation.
3. The term of this approval is for 3 years.
4. I wish you a successful and informative study.


Daryl Lindsay, Chair
University Advisory Committee
on Ethics in Behavioural Science Research

DL/bjk

Research Services, University of Saskatchewan
Kirk Hall Room 210, 117 Science Place, Saskatoon SK S7N 5C8 CANADA
Telephone: (306) 966-8576 Facsimile: (306) 966-8597

APPENDIX M

Screening Procedure Consent Form and Study Description

Consent Form

Dear Student:

You are being asked to participate in the first phase of a research study entitled "The effects of question construction on expository text comprehension" which is to be conducted by Dr. B.S. Randhawa and Thelma M. Gunn. The purpose of the study is to understand reading processes. Your task during this first phase is to complete a prior knowledge questionnaire concerning the heart. This information will be used to determine the number of individuals with high domain knowledge and low domain knowledge of the heart and its functions. On the basis of these findings, we hope to gather a population of eligible and willing participants for the second phase of the study who represent both knowledge levels. It is important to note that your name will be matched with an untraceable number to ensure privacy and anonymity. All results will be securely stored by Dr. Randhawa for a minimum of five years as required by University of Saskatchewan guidelines. The results will be reported in a doctoral dissertation and will likely be presented at professional academic conferences as well as published in the form of journal articles. Individual scores will not be reported as the primary purpose of the study is to understand the nature of the reading processes.

The duration of this phase of the study is approximately 10 minutes. If at any time you choose to withdraw from this project, you may do so without any consequences. In that case, all the data collected from you will be deleted from the study and it will be destroyed. Should you desire a summary of the results, please indicate below and give your address.

If you have any further questions regarding this project, please feel free to telephone either Dr. B.S. Randhawa (966-7661) or Thelma M. Gunn (966-7677).

Sincerely yours,

Thelma M. Gunn
Doctoral Student
Educational Psychology
University of Saskatchewan

I, _____, agree to participate in the above-mentioned research study. I understand that all information gathered is confidential and that my participation is voluntary. Moreover, I am allowed to withdraw from this study at any time without consequences. Should I choose to withdraw, all data collected on my behalf will be excluded from the study and it will be destroyed. I acknowledge that I have received a copy of this consent form.

Name

Date of consent

I want to receive a copy of the summary of the findings.

Yes

No

My address is: _____

Title: The Effects of Question Construction on Expository Text Comprehension

Researcher: Thelma M. Gunn

Supervisor: Dr. B.S. Randhawa

Professor, Department of Educational Psychology

Purpose for Study:

The ability to read and comprehend text is a well-recognized goal of instructional practice. It is crucial for success in our society since the written word is the primary vehicle for the transmission of ideas, thoughts, and understandings between individuals, societies, and eras. Considering the impact of expository text (i.e., informational textbooks, instructional manuals) on knowledge acquisition and utilization in the classroom and beyond, the need for research is evident. Unlike narrative text (i.e., literature) which is easier to understand and study, expository text is comprised of technical terms, unfamiliar theories, and prior knowledge; each of which the reader is typically lacking. Therefore, the purpose of this two phase study is to examine whether a specific reading comprehension technique (i.e., questioning) will not only enable students with high knowledge and low knowledge of text content better comprehend expository text, but help them to monitor their reading behaviours.

The purpose for the first phase of the study will be to determine those students with high domain knowledge and low domain knowledge of the heart. This is because the depth of one's prior knowledge strongly impacts the manner in which one reads a piece of text and the amount of information learned as a consequence. From the findings of the first phase, we hope to gather a population of eligible and willing participants for the second phase who represent both knowledge levels. During the second phase of the study, an equal number of high and low domain knowledge participants will be randomly assigned to three strategic questioning conditions: highly structured, semi-structured, and unstructured questions. Following a brief training period, they will be asked to read a few pages of scientific text. Some informal testing will occur immediately afterward. This will include a brief comprehension monitoring questionnaire, a 12-item posttest, a recall task, and a word association task. All instruction and testing will be conducted in groups of 10 to 20 participants.

Therefore, with the help of your participation, we hope to better understand how we may improve expository text comprehension for high domain knowledge and low domain knowledge students.

APPENDIX N

Experimental Procedure Consent Form and Study Description

Consent Form

Dear Student:

You are being asked to participate in the second phase of a research study entitled "The effects of question construction on expository text comprehension" which is to be conducted by Dr. B.S. Randhawa and Thelma M. Gunn. The purpose of the study is to understand reading processes. Your task during this second phase is to participate in some preliminary strategic training, to read a few pages of scientific text, and to engage in post-testing (i.e., word sorting tasks, a comprehension monitoring questionnaire, a recall task, and a post-test). It is important to note that your name will be matched with an untraceable number to ensure privacy and anonymity. All results will be securely stored by Dr. Randhawa for a minimum of five years as required by University of Saskatchewan guidelines. The results will be reported in a doctoral dissertation and will likely be presented at professional academic conferences as well as published in the form of journal articles. Individual scores will not be reported as the primary purpose of the study is to understand the nature of the reading processes.

The duration of this phase of the study is approximately 120 minutes. If at any time you choose to withdraw from this project, you may do so without any consequences. In that case, all the data collected from you will be deleted from the study and it will be destroyed. Should you desire a summary of the results, please indicate below and give your address.

If you have any further questions regarding this project, please feel free to telephone either Dr. B.S. Randhawa (966-7661) or Thelma M. Gunn (966-7677).

Sincerely yours,

Thelma M. Gunn
Doctoral Student
Educational Psychology
University of Saskatchewan

I, _____, agree to participate in the above-mentioned research study. I understand that all information gathered is confidential and that my participation is voluntary. Moreover, I am allowed to withdraw from this study at any time without consequences. Should I choose to withdraw, all data collected on my behalf will be excluded from the study and it will be destroyed. I acknowledge that I have received a copy of this consent form.

Name

Date of consent

I want to receive a copy of the summary of the findings.

Yes

No

My address is: _____

Title: The Effects of Question Construction on Expository Text Comprehension

Researcher: Thelma M. Gunn

Supervisor: Dr. B.S. Randhawa
Professor, Department of Educational Psychology

Purpose for Study:

The ability to read and comprehend text is a well-recognized goal of instructional practice. It is crucial for success in our society since the written word is the primary vehicle for the transmission of ideas, thoughts, and understandings between individuals, societies, and eras. Considering the impact of expository text (i.e., informational textbooks, instructional manuals) on knowledge acquisition and utilization in the classroom and beyond, the need for research is evident. Unlike narrative text (i.e., literature) which is easier to understand and study, expository text is comprised of technical terms, unfamiliar theories, and prior knowledge; each of which the reader is typically lacking. Therefore, the purpose of this two phase study is to examine whether a specific reading comprehension technique (i.e., questioning) will not only enable students with high knowledge and low knowledge of text content better comprehend expository text, but help them to monitor their reading behaviours.

The purpose for the first phase of the study was to determine those students with high domain knowledge and low domain knowledge of the heart. This is because the depth of one's prior knowledge strongly impacts the manner in which one reads a piece of text and the amount of information learned as a consequence. During the second phase of the study, an equal number of high and low domain knowledge participants have been randomly assigned to three strategic questioning conditions: highly structured, semi-structured, and unstructured questions. Following a brief training period, they will be asked to read a few pages of scientific text. Some informal testing will occur immediately afterward. This will include a brief comprehension monitoring questionnaire, a 12-item posttest, a recall task, and a word association task. All instruction and testing will be conducted in groups of 10 to 20 participants.

Therefore, with the help of your participation, we hope to better understand how we may improve expository text comprehension for high domain knowledge and low domain knowledge students.

APPENDIX O

Question Response Recording Sheets

Question and Answer Sheet

Identification # _____

QUESTION

ANSWER

QUESTION

ANSWER

QUESTION

ANSWER

QUESTION

ANSWER

QUESTION

ANSWER

QUESTION

ANSWER

APPENDIX P

**Correlation Matrix for Question Type
(Textbase, Elaborative Inference, Bridging Inference,
Problem Solving)**

Correlation Coefficients* for Question Types (Textbase, Elaborative Inference, Bridging Inference, and Problem-Solving)

	Textbase	Elaborative Inference	Bridging Inference	Problem Solving
Textbase	-	-	-	-
Elaborative Inference	.64	-	-	-
Bridging Inference	.67	.74	-	-
Problem-Solving	.44	.71	.67	-

*Correlations are significant at the .01 level (2-tailed).