READING COMPREHENSION ASSESSMENTS: EFFECT OF EPISTEMIC BELIEFS

ON TEXT AVAILABILITY AND QUESTION TYPE

by

Adrienne Farley Splinter

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Department of Educational Psychology

The University of Utah

August 2011

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of   **Adrienne Farley Splinter**

has been approved by the following supervisory committee members:

Anne Cook , Chair		12/15/2010
		Date Approved
Doug Hacker , Member		12/15/2010
		Date Approved
Dan Woltz , Member		12/15/2010
		Date Approved
Kristen Butcher , Member		12/15/2010
		Date Approved
Lauren Liang , Member		12/15/2010
		Date Approved

and by   **Elaine Clark**   , Chair of

the Department of   **Educational Psychology**

and by Charles A. Wight, Dean of The Graduate School.

ABSTRACT

This study used Kintsch's Construction Integration (CI) Model as a context for investigating reading comprehension assessment and its relationship to epistemic beliefs. Specifically, questions tied to levels of representation delineated in the CI model were used to investigate how individual differences may predict reading comprehension. Literal questions that tie to the textbase representation of text, and inference questions that tie to the situation model representation of text, were used to investigate these effects. In addition to question type, text availability was also manipulated in this study.

Previous studies have reported that a person's epistemic beliefs may have an effect on his or her ability to comprehend text. The current study was designed to investigate these findings using additional measures as covariates (i.e., vocabulary ability, background knowledge and working memory), and a test structure that is cognitively demanding (e.g., short answer questions to a variety of passages from three domains with and without text available). Participants completed individual measures and the reading assessment in a 2-hour session.

Although epistemic beliefs did account for a significant portion of variance in the reading assessment, this was very small, especially compared to the predictive validity of background knowledge and vocabulary ability. As predicted by the CI model, question

type was related to reading comprehension performance and this interacted with epistemic beliefs. Surprisingly, this effect was found with literal questions rather than inference questions. Text availability did not interact with epistemic beliefs. The results of this study suggest that when predicting reading comprehension, it is essential to use several individual differences variables, and that the relationship of epistemic beliefs with reading comprehension is less definitive than indicated in the research literature. Future work in reading comprehension research should establish other individual difference variables such as reader's interest level and specific strategy use when answering literal and inference questions using expository text.

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

CHAPTER I

INTRODUCTION

Reading comprehension is a multifaceted construct that involves several cognitive skills working together. Advances in studying the reading process have occurred over the past 2 decades, and our understanding has aided in the design of more reliable and valid measures of the construct. Although many of the influencing variables have been defined and studied, we still lack a complete account of the effects of individual differences on comprehension.

This study is designed to focus on specific individual difference variables and test elements in the hope of identifying and aligning a current theory of reading comprehension with comprehension assessments. Four individual difference variables will be considered within the context of a reading comprehension measure; however, only one of these is the main focus of the study. Specifically, this study proposes to investigate the interaction of readers' epistemic beliefs with test structure variables (availability of text and question type) on readers' ability to answer short answer questions after reading expository text. Before discussing how these specific test variables and individual difference variables may interact, I will present an overview of the theoretical issues underpinning this study, beginning with a model of reading comprehension, followed by a discussion of relevant individual difference variables.

Model of Reading Comprehension

Current reading theory suggests that individuals form multiple levels of representation of a text during reading. A reader has a lexical, textbase, and a situation model representation of the text (Kintsch, 1998; McNamara & Kintsch, 1996; Perffeti, 1997; van Dijk & Kintsch, 1983). The first level may be a surface level that allows the reader to recall verbatim what she read. Another level would allow a reader to paraphrase the gist of what was read, and still another would allow the reader to understand the message conveyed in the text and form a mental model (Kintsch, 1988, 1998). The latter two are the focus of this dissertation.

According to Kintsch's (1998) Construction Integration (CI) model, the reader establishes a textbase representation in which the words and sentences from a text are understood at their literal level. During the construction phase, a number of different cognitive processes take place. As a reader parses a sentence, a few propositions could be formed. For the following sentence:

> *Nearly all decisions involve trade-offs; there are advantages and disadvantages, costs and benefits, associated with every action every choice.*

One proposition may be that decisions have trade-offs. Another may be that there are consequences for decisions both positive and negative. Connections among propositions form a network, or series of networks, and these networks may involve connections that are direct or hierarchical in nature. As the construction phase continues, meanings of the words in propositions are activated and associations are made. Knowledge of the world as it relates to the propositions is activated, forming additional networks. It should be noted that this is an automatic process that can activate both relevant and irrelevant information (Myers & O'Brien, 1998; O'Brien & Myers, 1999). For the sample sentence described,

the reader may activate a recent decision she has made in the past that had an unexpected cost attached to it. During the integration phase, the reader uses these propositions to construct varying types of inferences about the text. She may, for example, infer from the semantic overlap of *costs* and *benefits* and *advantages* and *disadvantages* within this first sentence that the following text will emphasize the fact that all choices will have both good and bad aspects attached to them (Kintsch). As suggested by Mulligan (2006) and others, learning occurs when information from the text is integrated with what is already known (McNamara & Kintsch, 1996)

As reading continues, the propositions from subsequent construction cycles are mapped onto the previously established propositional network, and the reader begins to develop a situation model, or understanding, of the text. As the reader constructs a situation model, any new information has the potential for changing the established networks. During the integration process of reading, activation of previous knowledge and new information are linked together and may form more inferences. It may be difficult to determine what information in the mental model is a direct result of reading, and what was activated from general world knowledge in long term memory (Kintsch, 1998). Thus, the extent to which the reader establishes a mental model of the text, and can answer questions about the text, is influenced by the development of a complete and accurate textbase model in addition to constructing an accurate situation model. In addition, if reading goals are established that promote one model or the other, either model can be emphasized at the expense of the other (Kintsch). This could influence different levels of comprehension while reading; in other words, the reader may use different reading strategies based on goals. For purposes of studying the different levels

of cognitive processes readers use, researchers generally develop various reading comprehension assessments.

Using definitions derived from Kintsch (1998), the types of questions used in comprehension assessments are assumed to match the level of processing the reader has engaged in to understand the text. It is generally understood that literal questions can be answered based on readers' textbase representations, and inference questions require the situation model. The following example is used here to illustrate levels of questions:

> *If a firm purchases a new piece of equipment for $3000, there is an opportunity cost that $3000 could have been deposited in an interest earning account, or lent to another firm.*

An example of a literal level question, information that may be found in the textbase representation of the passage, is: "What is the opportunity cost for buying a new piece of equipment?" An inference question would require readers to make connections between their background knowledge and statements in the text to form an inference. The inference questions in this study are assumed to be tied to the situation model of text representation (McNamara & Kintsch, 1996; Perfetti, 1997). For the question, "Why would a firm want to purchase a new piece of equipment?" the reader would have to rely on their situation model of the text, because the answer to the question is not explicitly stated in the text. Inferences can be made when information in the text and activated general world knowledge is linked. For the purposes of this study, a question will be considered an inference question when the information needed to answer it is not explicitly stated in the passage, but can be logically inferred from the information presented.

Reading Comprehension and Individual Differences

Because of the complexity of reading comprehension, individual differences may affect any point in the construction and integration process. The primary focus of this study will be on individual differences in epistemic beliefs, although the influences of working memory, background knowledge, and vocabulary ability will also be considered.

Epistemic Beliefs

The main individual difference variable examined in this study is epistemic beliefs, or a person's general beliefs about knowledge. A reader with a more complex or relative view of knowledge may engage in a deeper search for meaning from the passage than a reader with a simplistic or concrete view of knowledge. As suggested by researchers using the CI Model, reader characteristics influence the level of text processing in which a reader engages, and thus the development of situation model representations (e.g., Mulligan, 2006). Mature epistemic beliefs about the complexity of learning may influence readers' abilities to use reading strategies effectively.

Epistemology has been discussed since the time of Plato. Questions about truth, what constitutes knowledge, how we know what to believe, and whether we can know truth have been discussed for centuries. Recently, it has been explored by psychologists trying to understand how beliefs may affect learning, thinking, and problem solving (Perry, 1970; Ryan, 1984; Schommer, 1990). Perry described the college students he studied as moving through a sequence of cognitive stages from a dualistic treatment of knowledge (i.e., right or wrong, true or false) to a more relativistic set of beliefs, where knowledge is an array of interpretation and integration of ideas. Ryan (1984) suggested that it is important to consider students' epistemic beliefs, "because a student's naïve

epistemology is likely to constrain the nature of the strategic hypotheses he or she develops and evaluates across a wide range of academic tasks (e.g., reading textbooks, taking lecture notes, answering examination questions, writing term papers)" (p. 1226). Educators and reading researchers also want to understand how epistemic beliefs influence students' ability to learn or interact with the cognitive processes involved in reading (Burton & Daneman, 2007; Kardash & Scholes, 1996; Rukavina & Daneman, 1996).

In order to understand why a measure of epistemic beliefs is necessary in reading comprehension assessment, it is helpful to realize that a person's ability to understand text may have a great deal to do with how they treat knowledge in general, especially new information in an unfamiliar domain. Schommer (1990) suggested that a student who believes that knowledge is simple, comprised of certain facts, and derived from authorities may oversimplify information found in text. This oversimplification may cause the reader to perform poorly in academic areas because he seeks to find simple answers when more complex solutions are more appropriate. Schommer argued that epistemic beliefs are more of a system of beliefs and that they are multidimensional and highly complex. Understanding these dimensions may help educators design their teaching practices to reach a broader spectrum of learners. In addition, there is significant debate over whether measures of epistemic beliefs can be considered general in nature, or whether one's beliefs may vary depending on the domain being considered (Bell & Linn, 2002; King & Kitchener, 2002). The argument would hold that a person's views of scientific knowledge may vary from his views in other areas such as humanities.

Some theorists suggest that epistemic beliefs are a component of metacognition, or an awareness of how we think (Pintrich, 2002). Other theorists argue that these measures (DeBacker et al., 2008) may be measuring levels of problem solving ability, or reading strategies. The definition of the construct varies depending on the theory. These arguments, both pro and con, emphasize the fact that epistemic belief measures, or similar metacognitive measures (i.e., Reflective Judgment Interview) may help to explain efforts made by skilled readers (King & Kitchener, 2002).

For the purposes of this study, epistemic beliefs are operationally defined by scores on the Schommer (1990) Epistemic Questionnaire (EQ) beliefs scale, a 63-item Likert-type instrument designed to measure five dimensions of beliefs about knowledge. It should be noted that an additional epistemic beliefs scale was added later in the study (Schraw, Bendixen & Dunkle, 2002)

These instruments presumably measure beliefs about knowledge with an acceptable amount of error. Participants respond to statements such as, "If you are going to be able to understand something, it will make sense to you the first time you hear it." using a scale of 1 (strongly disagree), to 5 (strongly agree). Analysis of these types of scales indicates that there are five dimensions of epistemic beliefs: (a) ability to learn is innate or improvable; (b) knowledge can be characterized as isolated pieces of information or an integration of concepts; (c) learning is quick or gradual; (d) knowledge is absolute or tentative; and (e) the source of knowledge ranges from omniscient authority or reason and empirical evidence (Perry, 1968; Ryan, 1984; Schommer-Aikins, 2004). Most researchers using these scales suggest that higher ratings indicate a more naïve epistemic belief (Burton & Daneman, 2007; Kardash & Scholes, 1996; Rukavina &

Daneman, 1996). The second and fourth factors have been used primarily by reading

researchers who have delineated them as knowledge is certain, absolute and simple, or

knowledge is relative, evolving and complex (Burton & Daneman, 2007; Kardash &

Scholes, 1996; Mason, Gava, & Boldrin, 2008; Rukavina & Daneman, 1996). With

regard to reading comprehension, beliefs about knowledge has implications of one's

ability to learn from text, or more specifically to abandon misconceptions in light of new

evidence presented in the text (Mason et al., 2008). Researchers using epistemic beliefs

scales generally use a median split cut-off score to determine mature versus naïve

epistemic beliefs held by a person. Researchers have noted that the scores represent a

continuous variable, and the cut-offs on such scales are not absolutes (Kardash &

Scholes, 1996). The current study analyzed the scores as a continuous variable. In

addition to Schommer's (1990) EQ scale, a second scale, the Epistemic Belief Inventory

(EBI), was added to this study to validate the construct of epistemic beliefs (Schraw et

al., 2002).

Several researchers have investigated the relationship of epistemic beliefs and

reading ability (Burton & Daneman, 2007; Kardash & Scholes, 1996; Mason, Gava, &

Boldrin, 2008; Rukavina & Daneman, 1996; Schommer, 1990; and Schommer-Aikins,

2004). In their study involving high school and college students exposed to two different

types of expository text with competing theories, Rukavina and Daneman (1996) found

that all students performed well on literal detail questions about two conflicting theories

involving the extinction of dinosaurs (see also Mason et al., 2008). However, those

students with more mature epistemic beliefs performed better on integrative short answer

and multiple choice questions than those with naïve epistemic beliefs, even when

controlling for working memory capacity. Naïve readers stated basic facts found in the text, whereas mature readers discussed comparative issues. These findings support the idea that readers with more mature epistemic beliefs were able to develop and access their situation model representations more effectively.

Previous research has generally focused on one topic or domain. For example, Kardash and Scholes (1996) investigated the effect of preexisting beliefs, epistemic beliefs, and need for cognition on readers' willingness to engage in effortful thinking about a controversial issue (e.g., Does HIV cause AIDS?). They found that people with naïve epistemic beliefs were less likely to engage in cognitively challenging tasks, such as reading about two sides of a controversial issue and interpreting who was right. Kardash and Scholes argued that those with more mature epistemic beliefs engaged in more effective reading comprehension strategies, which in turn made them better readers. The readers with mature epistemic beliefs responded with more tentative conclusions about the controversial issue of HIV causing AIDS, citing both sides of the issue. This suggests that certainty of knowledge, one of the dimensions represented in the epistemic belief scale, may be related to the depth of processing readers may engage in while reading, although this contention needs further investigation.

Burton and Daneman (2007) conducted another study of epistemic beliefs and reading, in which working memory and an epistemic beliefs scale were used to analyze accuracy on measures of reading. In order to control for background knowledge, the passages used were expository texts approximately 1,500 words in length based on familiar (diarrhea) and unfamiliar (fibromyalgia) diseases. Burton and Daneman used an on-line measure (eye-tracking), and an off-line measure (free recall) of what readers

remembered from the text. Using the on-line measure (i.e., regressions into target sentences), they found that epistemic beliefs interacted with working memory to influence readers' ability to acquire knowledge from text. Readers with low working memory capacity and mature epistemic beliefs reread text at strategic points, as compared to those with low working memory capacity and naïve epistemic beliefs. Burton and Daneman concluded that readers with more mature epistemic beliefs can compensate for their low working memory capacity by engaging in rereading at critical points in the text, particularly when background knowledge is low.

Research in discourse processing indicates that test takers with more mature epistemic beliefs may be able to compensate for limited background knowledge, vocabulary ability, and working memory, and they may be more successful in answering comprehension questions than readers with naïve epistemic beliefs (Burton & Daneman, 2007; Kardash & Scholes, 1996; Schommer, 1990). However, Schommer-Aikins (2004) cautioned that epistemic beliefs are embedded in other cognitive constructs, and she suggested that researchers must develop a deeper understanding of how to conceptualize these beliefs in a systemic model. For the purpose of this study, two epistemic scales were used to measure an individual difference variable that may help predict response patterns on reading comprehension tests (Schommer, 1990; Schraw et al., 2002). The focus initially was on the subscales of knowledge stability and knowledge complexity. As noted before, unlike previous research this study used the scale measures as a continuous variable.

It is important to note that concurrent validity research of epistemic belief scales have provided some contradictory evidence as to their internal consistency. In an analysis

of three epistemic belief scales, DeBacker et al. (2008) found that there was considerable variability in the reliability measures, and the operational definitions of the constructs these measures were designed to address. For example, for two of Schommer's (1990) original five factors, the original study reported internal consistencies around .79 (Learning is Quick), and .68 (Knowledge is Simple and Certain). However, DeBacker et al. were unable to replicate these reliability measures in their study. They reported measures of .40 (Knowledge is Quick), and .27 (Knowledge is Certain), with some subsets as low as .15. Therefore, they suggested that studies using these scales base their results on theoretical grounding rather than relying on the dimensionality ascribed to what the scales are measuring. Schommer's subscales are based on theoretical groupings prior to running a factor analysis, whereas the scale developed by Schraw et al. (2002) was based on a more traditional factor analysis.

The current research used Kintsch's (1998) CI model to delineate levels of comprehension. The two test structure variables (i.e., text availability and question type) were used as the basis for determining the level of reader's use of reading strategies (Mulligan, 2006; Rukavina & Daneman, 1996). Further, this study used the EQ and EBI scales as measures of a person's general beliefs about learning (i.e., that knowledge is complex, and therefore it takes effort to understand text from an unfamiliar domain). As suggested by previous research, there should be an interaction between a person's ability to answer inference questions and their beliefs about the complexity of learning. Schommer's (1990) epistemic beliefs scale was chosen for this study, because it is a more general measure of beliefs about knowledge, and has been used in several studies particularly in the realm of reading research using expository text (Burton & Daneman,

2007; Mason, Gava, & Boldrin, 2008; Rukavina & Daneman, 1996). The EBI scale (Schraw et al., 2002) provided additional validity to the construct being measured.

It is my hypothesis that readers with mature epistemic beliefs engage in deeper reading for understanding of the text because of their views of knowledge in general. This deeper reading is demonstrated by higher performance on the difficult task of generating short answers to inferential questions using expository text. In addition to using the scores from the epistemic beliefs scales as a continuous variable, the current study is designed to test the assumption that epistemic beliefs are more general in nature, rather than domain-specific. Therefore, I have included passages from three different domains (i.e., science, history, and business). Further, other individual difference variables (i.e., working memory, vocabulary ability, and background knowledge) were included as covariates in the analysis to evaluate epistemic beliefs' unique influence on reading comprehension.

Working Memory Capacity

As mentioned previously, Burton and Daneman (2007) found an interaction between working memory and epistemic beliefs, and they suggested that readers with mature epistemic beliefs could compensate for a low working memory span by developing better reading strategies. Researchers generally include a measure of working memory capacity because it is a good predictor of several cognitive functions, such as reading comprehension, vocabulary ability, and tests of reasoning (Daneman & Carpenter, 1980; Engle et al. 1999; Kyllonen & Christal, 1990; Radvansky, & Copeland, 2004). As suggested by Daneman and Carpenter (1980), there is strong support for investigating the relationship between working memory and reading comprehension.

Larger working memory capacity has a positive influence on reading comprehension. Information becomes part of working memory through various routes such as perceptual input from the text, and activation of knowledge retrieved from long term memory. With respect to the CI model, working memory efficiency is essential to processing the propositional networks that form during the cycles of construction and integration of new text (Kintsch, 1998). Working memory capacity may influence the amount of information readers can activate and/or integrate during comprehension construction cycles. This could hinder a reader's ability to develop accurate situation model representations, thus affecting his or her ability to respond to questions about text, particularly inferential questions. For example, Daneman and Hannon (2001) found that readers with high working memory capacity have higher accuracy rates on multiple choice reading tests than those with low working memory capacity. Further, Perfetti and Lesgold (1977) suggested that better readers may be those who use working memory more efficiently.

Several working memory measures have been developed over the past few decades. The measure generally used by reading researchers is Daneman and Carpenter's (1980) reading span measure. The reading span measure presents a series of sentences to the participants and then they are given an intervening task. Participants are then asked to recall the last word of each sentence when prompted over the course of three trials, and these trials range from two to six sentences each. The score is determined by the number of words participants remember across all trials. The reading span measure was designed to tax both the processing and storage functions of working memory. To validate their measure, Daneman and Carpenter correlated scores on their measure with a researcher-designed reading comprehension measure of fact questions, pronominal reference

questions, and the Verbal SAT. Those correlations were .72, .90, and .59, respectively.

Another working memory measure (word span) yielded correlations of .37, .33, and .35,

respectively. The two span tests were moderately correlated at .55. A substantial body of

evidence using Daneman and Carpenter's measure demonstrates that it is a good

predictor of reading comprehension (For a review of studies see Daneman & Merikle,

1996).

 However, Engle et al. (1999) argued that working memory measures reflect

differences in controlled attention capability and the capacity to maintain a level of

activation for designated items while interference from other activation is competing for

attention. Engle et al. suggested that the choice of task should be guided by logic and

previous empirical research. Their study indicated that span tasks correlated slightly

higher with verbal measures such as the Verbal Scholastic Aptitude Test (VSAT) (.49),

than with a measure of attention (.39). In support of this idea, other researchers have

suggested that the reading span test (RST) should not be used as a single measure, but

should be understood as two factors (Whitney, Arnett, Driver, & Budd, 2001). One factor

is processing speed; the other factor is susceptibility to interference.

The implication for this with regard to reading comprehension is that low span

readers could have problems with forming inferences because of semantic relationships

or syntactic structures, thus decreasing processing speed. They could also have a problem

with an inability to disregard irrelevant information (susceptibility to interference). If the

measure is combined into a single score, this difference would not be reflected.

Therefore, working memory tasks designed to measure attention may be more suitable

for the current study, because they do not correlate as highly with reading comprehension

and vocabulary measures, and have been used in previous research to demonstrate variability in working memory capacity (Kyllomen & Christal, 1990; Radvansky & Copeland, 2004).

The ABCD working memory task presents three sets of relationships for four letters (ABCD). The computer screen presents set 1 (A→B), then set 2 (C → D), then which set comes first (set 1 → set 2). Participants then select a correct answer from a multiple choice list. Order and negation are counterbalanced over 32 trials.

For the present study, the ABCD working memory task was selected because it is an attention-based working memory task that does not include linguistic elements such as word meanings. Was and Woltz (2007) found low correlations with the ABCD working memory task and a verbal ability measure (r =.32). Further, in a study designed to analyze the relationship between reasoning and working memory capacity, Kyllonen and Christal (1990) reported that a similar ABCD working memory task correlated with general word knowledge, general science, and paragraph comprehension tests at .18, .18, and .17, respectively. However, it did load on a working memory factor at .44. Kyllonen and Christal argued that it is measuring the working memory construct, but may not correlate as high with measures involving semantic information (i.e., Daneman & Carpenter, 1980). It should be noted, however, that the ABCD measure was the lowest of the four other comparative working memory tasks, but their loadings on the working memory measure were not significantly different: AB general (.53), Digit Span (.53), AB assignment (.75), and mental arithmetic (.56).

Further support for using the ABCD working memory task in this study can be found in Ackerman, Beier, and Boyle's (2002) study analyzing the underlying construct

of working memory. They reported a correlation of .63 for a similar task in a confirmatory factor analysis for working memory ability. They also provided further evidence that the ABCD task correlated only slightly with vocabulary ability (.21), and the Nelson-Denny reading comprehension test (.35).

Background Knowledge

In addition to epistemic beliefs and working memory, another predictor of reading comprehension ability is background knowledge. Studies show that readers' performance on comprehension tests depends on their knowledge of the topic being presented in the texts (Afflerbach, 1986; Chi, Feltovitch, & Glaser, 1981; Cóte, Goldman, & Saul, 1998; Lundeberg, 1987; McNamara & Kintsch, 1996; McNamara & McDaniel, 2004; Mulligan, 2006). It should be noted that this is particularly true of expository text; this aspect will be considered in greater detail later in this dissertation.

Research has found that domain knowledge is a high predictor of readers' ability to answer comprehension questions, both literal and inference, but there is evidence that this occurs more often as readers age and mature. In their study using think aloud protocols, Cóte et al. (1998) argued that prior knowledge aided sixth graders' reading of informational text, but these readers did not strategically use prior knowledge to build their situation models of the text. This was demonstrated by good local cohesion (textbase) during the think aloud, but it did not transfer to postreading questioning of global coherence (situation model). For example, when they read a confusing statement, readers engaged in self-explanations as to what it might mean, but when answering comprehension questions they failed to integrate these explanations. In reference to the

CI model, this would suggest that readers are forming accurate textbase representations, but fail to develop accurate situation model representations of the text (Kintsch, 1998).

The findings of Cóte et al. (1998) supported a previous study conducted by Scardamalia and Bereiter, in which global coherence was not present in seventh- and eighth-grade readers, but was present in 10th-grade readers. This may indicate that background knowledge facilitates the readers' ability to make inferences, but may vary between individuals and maturity levels. This suggests a strong association between maturity level and ability to use background knowledge in developing good situation models, which may tie to a person's beliefs about the complexity of learning and structure of knowledge, as Schommer (1990) demonstrated.

Background knowledge effects on the development of situation model representation may be dependent on other variables such as test structure. For example, Mulligan (2006) found that background knowledge's effect on reading comprehension differed based on the type of questions asked. Readers were more accurate on literal question recall than on forming accurate responses to inference questions. She concluded that, in many cases, participants who had specific prior knowledge necessary to form situation model representation failed to link that knowledge to the text and generate essay responses to inference questions. It may be that readers with mature epistemic beliefs use strategies designed to link prior knowledge with current text, although Mulligan did not address this issue specifically.

Test designers must control for the effects of background knowledge on reading comprehension assessments. With regard to correcting for domain knowledge, some researchers give a pretest to assess how much background knowledge the participants

have in a specific subject (Mulligan, 2006). The pretest could include having the participants write short essays, or take a multiple choice test. A potential problem with this method is that the pretest could prime prior knowledge, which could then under- or overestimate participants' knowledge. The pretest also could set up expectations for the reading test where readers use information from the pretest to develop strategies for taking the test. As stated previously in the CI model, if goals are established a reader may emphasize developing textbase over the situation model representation when reading (Kintsch, 1998).

An alternative to a pretest is to embed the questions measuring background knowledge in the reading comprehension test. The method adopted by this study has been used with various studies of both listening and reading comprehension measures (Burton & Daneman, 2007; McNamara, Kintsch, Songer, & Kintsch, 1996; Shapiro, 2004; Voss & Silfies, 1996; Was & Woltz, 2007). For example, Was and Woltz questioned readers about background knowledge using content not covered in the passage presented, but taken from the same source as the passage. For example, if the source of the passage is taken from an economics textbook, the background knowledge questions would be developed from the same chapter, but not the portion of the passage used in the actual test. The background knowledge questions are presented with the other test questions in the instrument. It stands to reason that readers with low background knowledge in the subject will not be able to activate prior knowledge to answer these questions. Thus, as Mulligan (2006) found, readers with more domain knowledge may be more accurate with literal versus inference questions, especially when the text is available during question answering.

Vocabulary Ability

In addition to epistemic beliefs, working memory, and background knowledge, vocabulary ability has a very high correlation with reading comprehension ability. Pearson, Hiebert, and Kamil (2007) suggested that general vocabulary can be defined in many ways. In essence, it is the sum of words employed by a language, group or individual, or work in a field of knowledge. As stated by the National Reading Panel (NRP: National Institute of Child, Health and Human Development [NIHCD], 2000), vocabulary has long been viewed as an important factor in reading achievement. Even so, researchers admit that it presents problems with assessment of reading comprehension. The NRP treated it as a separate attribute and analyzed it as a separate component from reading comprehension. Although some researchers suggest that this ability is really what is being measured by most multiple choice reading tests (Valencia & Pearson, 1987), the NRP countered this argument stating that although vocabulary ability and reading comprehension are two separate "skills" they remain strongly related. Early researchers measured two factors in comprehension assessments: vocabulary and "gist" (Davis, 1942). Others stated that there is "something independent about vocabulary" that could be measured separately (e.g., Pearson & Hamm, 2005). They suggested that as readers' vocabulary ability develops, so too does their background knowledge and this in turn increases their ability to comprehend more complex text. This may allow readers to change their concepts regarding knowledge.

The debate over vocabulary demonstrates the complexity of the reading comprehension issue, because even in a fairly homogenous group there may be a high degree of variability in word knowledge. Many reading comprehension instruments build

the vocabulary assessment within the design of the test (Everson, Osterlind, Dogan, & Tirre, 2007). For example, in a practice item for the SAT Critical Reading Assessment, questions are structured with the vocabulary assessment embedded within the comprehension sections of the test. For a passage about a Chinese paleontology student discovering a fossil, the text reads:

> *The rock was still wet. The animal was glistening, like it was still swimming," recalls Hou Xianguang. Hou discovered the unusual fossil while surveying rocks as a paleontology graduate student in 1984, near the Chinese town of Chengjiang. "My teachers always talked about the Burgess Shale. . .*

A test item is presented:

> In line 5, "surveying" most nearly means:
>   (a) calculating the value of
>   (b) examining comprehensively
>   (c) determining the boundaries of
>   (d) polling randomly
>   (e) conducting statistical study of

This type of vocabulary assessment has an advantage of allowing the vocabulary to be tested within context. However, there is a disadvantage because there is an overlap of vocabulary and comprehension within the same measure, but they are reported as separate scores. This is problematic because one's skills in one area may compensate for lack of ability in the other area, and this would not be evident in the scores reported. The NRP recognized this problem and suggested that although tests report these two scores as separate, they may be measuring a combined construct.

Although Pearson, Hiebert, and Kamil (2007) tentatively endorsed the movement of the National Assessment of Educational Progress (NAEP) toward integrating vocabulary assessment within the comprehension portion of their tests, they cautioned that these new assessments have not been fully tested for construct validity, especially

given the new emphasis using expository text. This issue will be discussed in more detail in the section on discourse type. As Pearson et al. suggested, "[T]he goal (of the NAEP) is to report vocabulary separately, assuming that the construct, as measured, stands up to the psychometric validation of its statistical independence. . ." (p. 287). Currently, the NAEP is conducting several validation studies to determine whether the new design is reliable and valid.

Because vocabulary, like working memory, has been demonstrated to be a separate albeit related construct to reading, it is important to assess it independently from reading comprehension. Therefore, the use of a separate instrument was selected to determine vocabulary ability's unique contribution to reading comprehension.

<div align="center">Reading Comprehension Test Structure</div>

As suggested in the previous sections, factors such as working memory, background knowledge, vocabulary ability, and epistemic beliefs may interact with test structure in varying ways depending on the design of the assessment. Recently, several researchers have examined whether current methods of assessment are valid in light of the new theories in reading processes (for a discussion see Afflerbach, 2008). As Cain, Oakhill, and Bryant (2004) have argued, reading assessments tend to be driven by psychometric properties rather than theoretical models of reading. In addition, the current emphasis on educational accountability may promote teaching to the test rather than focusing on developing higher level reading skills necessary for the social demands of an information laden society. As Afflerbach (2008) suggested, "advances in theoretical and practical knowledge require that high quality reading assessments correspond to current knowledge of the content and processes that should be assessed in reading" (p. 152).

Current measurements of reading comprehension report a single score. Does this score represent the same construct with every instrument? Mulligan (2006) suggested that it does not. There are several factors within the measures that vary considerably from one another (e.g., type of text used and type of questions asked). Therefore, it is essential to study specific facets of a test's structure when determining the interaction of individual differences and the comprehension measures. The current study is investigating the impacts of specific test characteristics (i.e., text availability and question type) on reading comprehension. It is designed to investigate readers' ability to answer literal and inference questions using a short answer format, with expository text. These facets were selected due to the difficulty readers have with this particular test design. In addition, this is a practice similar to what is required of readers in real world situations (i.e., general college courses). Professors generally assign reading from textbooks then test on the material covered in the readings. Students must respond to questions about the information without having the source material available.

Typical reading comprehension tests have many similar structural characteristics. For example, these tests typically involve a brief passage of approximately 100 to 300 words (from either expository or narrative text), followed by multiple choice or short answer questions written at varying levels of difficulty. The very nature and purpose of standardized reading assessments, hereafter referred to as "test(s)," are usually to rank students on reading ability (Bennet, 1998). Although the tests have gone through changes over the last 4 decades, essentially they have returned to the structure used in the 1970s (Sarroub & Pearson, 1998). Assessment includes narrative and nonnarrative passages with both literal and inference, with the majority of the questions being literal.

Given the move toward national standards of student achievement (NCLB), reading assessments have been used to determine students' educational achievement. Although new national assessments are being developed, most state tests rely on the structure described previously. For example, Utah uses the Utah Basic Skills Competency Test (USOE, 2002), which consists of eight passages including package labels, bus schedules, journal articles, and short stories. Each passage is followed by five to eight multiple choice questions. Approximately 80% of the questions are literal. The results from the test are currently used to determine if students will receive their high school diplomas. One score is provided that ranks the students on reading competency. At issue with tests of this nature is that they do not take into consideration the complexity of reading comprehension, and there is further risk that instruction in reading may focus on reading for literal representation, or the textbase level, rather than a deeper level of comprehension or the situation model (Kintsch, 1998). This study is designed to investigate specific facets of test structure deemed as the most difficult, and provide support for alignment of tests to reading comprehension models.

Studies concerning the validity of reading assessments have focused on differences in: availability of text during question answering (Mulligan, 2006; Ozuru et al. 2007); levels of questions, such as literal versus inferential (Magliano & Radvansky, 2001; Singer, 1994; Singer & O'Connell, 2003; Wiley & Myers, 2003); types of question format, such as multiple choice versus short answer (Campbell, 1999); and types of discourse, such as expository versus narrative (Mulligan, 2006; Rukavina & Daneman, 1996). Although these test factors are related, they have generally been studied in isolation in reading research literature. This study is designed to integrate several of

these factors concurrently to determine how test factors influence readers' ability to access their textbase and situation model representations as presented in the CI model (Kintsch, 1998).

<u>Text Availability</u>

Text availability during question answering is perhaps one of the most important and least studied variables in reading measurements. When an individual reads a passage and is required to remember and respond to a question without the text available, she may have to engage in deeper reading. It has been suggested that literal questions are easier to respond to than inference questions, and by requiring responses without text available this difficulty level may be more pronounced (Mulligan, 2006). Therefore, if tests are designed to measure reading comprehension at various levels, text availability is an important factor to consider. It is generally understood that readers' ability to answer questions without the text available measures their ability to access their mental model of what they read (Anderson, 1978; Kintsch, 1998). It is further assumed that the textbase model is a more surface structure of text representation than the situation model. Therefore, readers may be able to generate responses to literal questions, even without the text available, more easily than they can generate responses to inference questions based on their situation model. Further, when the text is available during question answering this may only be measuring the readers' ability to process the text on-line. For example, readers could look at the question first, and then skim the passage for a response.

Studies have shown that not having the text available during question answering decreases performance on reading comprehension assessments (Katz et al., 1990;

Mulligan, 2006; Ozuru et al., 2007). Mulligan's study involved college students reading six texts taken from the SAT practice tests and college textbooks. These were expository texts that varied in length from 200 to 700 words. She had participants read the text and then write a short retelling of the passage. She referred to this as memory (M) for textbase representation based on Kintsch's (1998) CI model. Following this, the participants were asked to write a response to an inference (I) question, which aligns with readers' situation model representations. Her results showed that M scores were higher overall than the I scores, and in the without text available condition there were lower scores in general. She found that question type differed as a function of text availability. She further noted that although readers were able to construct good memory essays for the text, this did not guarantee that they could also generate good inference responses. She stated that the M score with text available was an upper bound for what readers remember, but the I score, which represents the situation model, varies a great deal depending on individual differences as measured by prior knowledge and interest.

Ozuru et al. (2007) examined three independent variables (format of questions, text availability, and question difficulty level) using expository text. They found inconsistent patterns of results. For example, in their first experiment, there were high correlations between multiple choice and short answer for literal questions, but not for inference questions. In a second experiment to test differences in levels of question and text availability while accounting for background knowledge, their analysis indicated it was very difficult for readers to respond to open-ended inference questions when the text was not available. Ozuru et al. suggested that when the text was available readers relied on rereading of the text to generate answers. The study demonstrated the difficulty in

measuring the complex aspects of reading comprehension. It further suggests that text availability could be a strong factor in determining readers' level of text comprehension.

The findings of both Mulligan (2006) and Ozuru et al. (2007) support the idea that without the text being available, readers must rely on their situation model to answer inference questions, and that requiring short answer responses is much more difficult for readers. Both of these studies support the idea that when text is made available readers tend to rely on rereading portions of the text or skimming for answers. This is made even more difficult when requiring readers to generate short answer questions using expository text. As suggested by Mulligan, individual differences may mitigate the difficulty level, but it is difficult to determine which factors influence readers the most. Therefore, text availability is manipulated in the present study to determine if epistemic beliefs influence literal and inferential questions differently.

<u>Types of Questions</u>

According to the CI model, different questions are thought to tap into varying levels of cognitive processing. For testing purposes, the levels of questions are used to determine readers' ability to read for understanding at the surface level of the text (textbase), and also at a deeper cognitive level or situation model (Kintsch, 1998). Most tests have a higher number of literal questions (e.g., Gates MacGinite Reading Tests: MacGinite, MacGinite, Maria, & Dreyer, 2002; Nelson-Denny Reading Comprehension Assessment: Brown, Fishco, & Hanna, 1993) than inference questions; however, current trends of test construction are moving toward increasing the number of inference questions to assess the mental model the reader has constructed (Embretson & Wetzel, 1987; Gorin, 2005; Gorin, Embretson, & Sheehan, 2002; Sheehan & Ginther, 2001).

Including inference questions is essential in tests to measure level of understanding. A great deal of reading that individuals engage in every day involves expository text. But including inference questions is problematic with expository genre when testing is outside the context of a classroom, where teachers are building on the students' knowledge and can address confusions. Perhaps this is the reason most tests require test takers to answer few inference questions. Analyses of current tests demonstrated that a large percentage of the items were found to be "trivial" or peripheral in nature (Johnston & Afflerbach, 1982); that is, the questions tapped the textbase representation, rather than the situation model. Critics of this practice point to the adverse effect this may have on students. Students may come to rely on surface processing of text, because that is how reading is assessed. Further, teachers may feel obligated to teach to the test rather than how to read for deeper understanding. This may have an adverse effect on education in general. Smith (1991) suggested, "each day spent teaching to a test that represents only a narrow sample of what reading is can contribute to decreased teacher motivation and enthusiasm" (p. 156).

Answering inference questions in reading assessments using expository text can be difficult. Studies have found that there is a difference in the likelihood of readers making inferences when reading narratives, as opposed to expository text, and in familiar and unfamiliar domains. For example, in Bowyer-Crane and Snowling's (2005) study of poor versus normal comprehenders, the former were much less likely to use background knowledge and lexical cues to form different types of inferences (i.e., elaborative, cohesive, knowledge- based) when reading expository texts, but they responded in the normal range when tested with predominantly fictional text. This suggests that

comprehension tests may yield different results due to the types of questions asked, and text type used.

Other researchers support this argument by demonstrating that readers are much more likely to draw inferences when reading narrative text (Magliano & Radvansky, 2001), because they can draw on general world knowledge about characters and their motives. Readers who are unfamiliar with a subject presented in expository text, on the other hand, make fewer inferences, unless directed to do so with a prompt (Kintsch, 1998). As mentioned previously, the reader's background knowledge has a differentiating influence on inference generation (Mulligan, 2006), particularly with expository text.

As noted previously, in the study conducted by Ozuru et al. (2007) literal questions were easier to answer than inferential questions when participants were unfamiliar with a topic; however, high background knowledge presented conflicting results in their study. The researchers suggested that at times background knowledge interferes with answering literal questions to a greater extent than inference questions, but their findings were inconclusive. Thus, further investigation of question type using expository text is needed.

Support for the difficulty of assessing expository text with different question types can also be found in the discourse processing literature. Singer (1988) stated that bridging causal inferences must be validated with reference to general world knowledge. Although his studies focused on narratives and are tested with inconsistency paradigms, others have used his definition with expository text (Wiley & Myers, 2003). Wiley and Myers defined bridging causal inferences as those that are derived from two separate sections of the text that may or may not be resolved with general world knowledge. In

their study, they presented readers with science texts that had two separate propositions followed by a statement with either a consistent and inconsistent inference. They monitored reading time on the target sentence to determine if the inferences had been made. They concluded that if both propositions were in close proximity to one another in the text, there were longer reading times on an inconsistent inference. If the propositions were separated by a filler passage, there was not a slow down on the inconsistent inferences. They argued that readers may be able to make bridging causal inferences, but that these may not be as robust when propositions are not stated in close proximity, particularly with expository text.

In the current study, readers were required to rely on the situation model representations they had established while reading in order to answer inferential questions particularly when text was not available. According to Kintsch's (1998) CI model, readers' general world knowledge is activated along with the text presented. It is assumed that the level of cognitive processing matches the degree of difficulty of the questions. Therefore, the literal questions should be easier than the inference questions, but this may be mitigated by factors such as text availability, working memory, vocabulary ability, background knowledge and epistemic beliefs.

Discourse Type

The issue of how best to assess the situation model so it matches the real world requirements of expository reading is very complex. Tests that use expository passages which require readers to use their background knowledge in connection with discourse may provide better evidence of the situation model. For instance, expository passages

taken from textbooks may provide several layers of propositions that are well developed, and readers can form a substantial situation model of the text.

Deane et al. (2006) suggested that there are several issues related to discourse length and type that can create difficulty for readers. These include narrative and nonnarrative discourse, vocabulary, and structures of academic versus nonacademic discourse. One of the major issues with assessing readers with nonnarrative academic discourse is requiring the readers to make inferences beyond the text. Thus, great care must be taken to develop tests that are accurate in measuring this ability.

The present study is using the guidelines established by the NAEP, which defines expository texts as including informational trade books, textbooks, news articles, feature articles, encyclopedic entries, and historical documents (Moss, 2008). As Moss suggested, the NAEP council has recommended an increase in inclusion of more informational text in literacy instruction, starting at 50% for fourth grade and gradually increasing to 70% for 12th grade. This is in part to address the fourth-grade slump exhibited historically with nationwide reading assessments.

Moss (2008) suggested that the drop in proficiency scores in fourth grade is due to the inclusion of expository text included in the testing instrument of the NAEP. She conducted a study to analyze the basal readers adopted by California in 2006 to determine if they followed the NAEP guidelines for inclusion of expository text. She found that the basal readers are including more nonnarrative texts, but only half of the texts were expository in nature. Further, she found that most of the texts were biographical. This overrepresentation of a specific genre should be addressed by national and state education regulatory boards. A wide array of expository text should be provided to teach reading

instruction. From an ecological validity standpoint, tests should include textbook type passages from subjects studied by most students in schools (i.e., history, business, and science)

The overrepresentation of biographical genre in tests can be further illustrated by other standardized reading tests. For example, in the Nelson Denny Form G (Brown, Fishco, & Hanna, 1993) there are six passages with approximately 200 words per passage. Three passages are biographical, and three are informational, and each selection has five multiple choice questions primarily asking literal questions. Standardized tests use short test-designer created passages to increase the reliability of the test, but critics argue that this practice may undermine the test's validity (Valencia & Pearson, 1987). For example, on the SAT, there are four passages and 40 multiple choice questions. The discourse types are essays, science journals, history articles, etc. This is a timed test; therefore, individuals read the passage then answer approximately 15 questions in 15 minutes (Daneman & Hannon, 2001). Using shorter passages facilitates presenting more items to improve test reliability, but some critics suggest these shortened paragraphs lack coherence, and are not necessarily authentic reading (Duran, McCarthy, Graesser, & McNamara, 2007). Duran et al. suggested that reading measures should use excerpts from a section of text that has a high degree of coherent propositions, and readers can generate inferences based on the information in the text.

As mentioned previously, it may be difficult for readers to generate causal inferences when reading expository text as opposed to narratives, especially when they are unfamiliar with the subject (Singer & O'Connell, 2003). This may be due to several factors; for example, readers can draw on more general background knowledge, such as

their scripts or schemas, when they read narratives (Magliano, & Millis, 2003; Millis, Magliano & Todaro, 2006). These may not be available when reading expository texts such as excerpts from science or business textbooks. Without strong background knowledge about a subject, readers must rely heavily on the text to build their situation models. There may also be preconceived notions (even misconceptions) regarding the subject matter that may interfere with reading comprehension (Rukavina & Daneman, 1996).

The current study is designed to explore the ability of readers to answer literal and inference questions about a variety of expository texts from three different domains (i.e., business, history, and science), using the main discourse type found in educational settings (textbooks). In addition to text availability, level of question, and type of discourse used, the format of the question also affects readers' performance on tests.

Question Format

Researchers have found variability in test takers' performance on types of question formats used in tests (Campbell, 1999; Manhart, 1996). Campbell conducted a think aloud protocol with eighth graders using multiple choice and short answer formats with the NAEP test. He found that readers differed in their thinking processes as a function of question format. For example, readers engaged in higher level cognitive processes in different ways. For multiple choice questions, readers tended to use the available answers to frame their responses. While answering the short answer questions, readers tended to construct their responses based on what they had talked about as they were reading. Campbell (1999) recommended using both types of format because of the different processes the formats elicit. As noted previously, recent cognitive research in

assessment has shown differences in performance with question format, but this had different results when the text was made available (Mulligan, 2006; Ozuru et al., 2007).

In support of the notion that the two formats represent two different cognitive processes, Manhart (1996) conducted a factor analysis for determining whether multiple choice and short answer tests measure the same construct. Whereas Campbell's (1999) study used literary texts, Manhart's study used the Tests of Achievement and Proficiency (TAP) performance assessment for science, or expository text. In Manhart's study of high school students, he suggested that short answer and multiple choice questions measured two different constructs. Although he does not explicitly state what the cognitive processes are, he suggests that recognition on multiple choice formats involved a more superficial level of processing than recall.

The current study is designed to investigate comprehension ability with short answer format, because this format is considered to be more difficult than multiple choice format, because it requires a generated response to both literal and inference questions (Kintsch, 1998). The difficulty of the short answer format, particularly with the inference questions in the no text available condition, should provide a strong test of whether readers with mature epistemic beliefs have engaged in a deeper understanding of the texts (Kintsch, 1998). Further, the influence of reading strategies developed by readers with mature epistemic beliefs may facilitate accurate responses to literal questions as well.

<div align="center">Statement of Problem</div>

In this experiment, readers were presented with expository text passages and were asked short answer questions either with or without the text available during question answering. Readers were given both literal and inference questions to assess

understanding at the textbase level and the situation model of text representation, respectively (Kintsch, 1998). In addition, several individual difference variables (i.e., epistemic beliefs, working memory, background knowledge, and vocabulary ability) were assessed. The questions being addressed were:

- After factoring out other individual differences, do epistemic beliefs (EB) and text availability (TA) interact to predict a person's accuracy on reading comprehension assessments?

- Do epistemic beliefs (EB) and question type (QT) interact?

- Is there an interaction between QT and TA?

For the purposes of controlling for individual differences, separate measures of working memory (WM), background knowledge (BK), and vocabulary ability (VA) were included in the analyses.

Hypotheses and Research Design

The present study attempted to determine if readers with mature epistemic beliefs are more accurate in answering both literal and inferential questions than those with naïve epistemic beliefs when the text is not available during question answering. This is based on the assumption that readers with mature epistemic beliefs will have processed the text at a deeper level than naïve participants, thus producing richer textbase and situation model representations in memory. As a result, mature readers are able to answer both literal and inference questions accurately. Naïve readers, on the other hand, may only be able to develop accurate textbase and situation model representations when the text is available. The questions act as prompts, and when the text is available readers may generate accurate answers. Hence, there should be very little difference between mature

and naïve readers when the text is available, but the differences in question answering ability should be more apparent when the text is not available.

In the current study, every participant performed a reading comprehension task under four treatment conditions (i.e., with and without text available, and literal and inference questions). An analysis was then performed to test whether epistemic beliefs, in addition to the other individual differences variables mentioned previously, interacted with the treatment conditions as predicted. A repeated measure Analysis of Covariance (ANCOVA) was selected to test for main effects of text availability and question type, and interactions with epistemic beliefs and other individual difference measures (i.e., working memory, background knowledge, and vocabulary ability). The latter measures were considered covariates in the design. Time measures were also recorded for the participants. Participants were told whether the text was going to be available during question answering. It was predicted that there would be a different pattern of time measures as a function of text availability. An important aspect of reading comprehension measures used in education settings is that most are timed. However, for the purpose of this study, no time limit was imposed during the reading comprehension portion. Primarily, time limits were used to test for adherence to testing protocol. The time measures were not a specific focus in this study, but may help support the assumption that mature readers take longer to read passages when they know the text will not be available during question answering.

CHAPTER II

METHODS

Participants

One hundred and fifty-seven participants were recruited from the Educational Psychology subject pool at the University of Utah, and from other university courses. Demographic information is presented in Table 1. Most students were undergraduates in their third or fourth year of college with a median age of 25; however, some were older graduate students. Participants received partial course credit for their participation in the research. There were approximately 70% females, and the majority of the participants were Caucasian. In addition to the experimental participants, there were approximately 60 participants involved in a pilot study conducted to norm the materials.

The participants were generally education majors. They signed up for a 2-hour session using an on-line recruitment website. Each participant was randomly assigned to a version of the study when they reported to the research lab located on campus. One to seven participants were run at a time; however, there were separate carrels for each participant, and headphones were available to block out distractions and noise.

Table 1.

Participant Demographic Information: Gender, Primary Language, Age, and Class

Variable	Category	Percentage
Gender	Female	70
English as First Language		99
Age	18-21	40
	22-25	25
	26-29	16
	30-35	11
	36 and over	8
Class	Freshman	6
	Sophomore	17
	Junior	34
	Senior	28
	Graduate	13

<div style="text-align:center"><u>Apparatus</u></div>

Participants performed measures on PCs with SVGA monitors and standard keyboards. Programming and administration of all tasks was provided by E-Prime software (Schneider, Eschman, & Zuccolotto, 2002). E-Prime controlled the stimulus, presentation, and data collection except as noted for the different measures.

<div style="text-align:center"><u>Materials</u></div>

<u>Epistemic Beliefs Measures</u>

As described earlier, epistemic beliefs measures are self-report, using a Likert-type scale format. For example in the scales used in this study (Schommer, 1990; Schraw et al., 2002), participants responded to statements such as, "If scientists try hard enough they can find the truth about anything." The statements were randomly presented on the computer screen, and participants responded via number pad key presses (1 for strongly disagree to 5 for strongly agree). Statements were only presented once, and participants self-paced through the series, but they could not return to previous statements. There were a total of 84 items in the set. The specific statements used in the study, along with findings, are presented in the Results section. There are generally two factors reading comprehension researchers use from this scale. They are "stability of knowledge" and "complexity of learning." Previous studies using these scales reported varying reliability quotients of .30 to .80 (DeBacker et al., 2008; Schommer, 1990). Exploratory factor analyses of the two scales were conducted. Reliability of internal consistency was also determined, along with intercorrelations with the other individual measures. Selection of the Factor used in this study will be discussed in the Results section.

Working Memory Measure

Variants of the ABCD working memory measure chosen for this study have been used with previous studies (Ackerman, Beier, & Boyle, 2002; Kyllonen & Christal, 1990; Was & Woltz, 2007; Woltz, 1988). These measures correlate at a moderate level with other working memory (WM) measures. The measure is described as an attention task rather than a span task. The ABCD WM measure required participants to interpret three graphically presented statements that together defined the order of the letters A, B, C, and D. One symbol defined the order of A and B (A→B; interpreted AB). Another symbol defined the order of C and D (D ← C; interpreted as CD). The third symbol defined the order of AB relative to CD (e.g., Set 1 ← Set 2; interpreted as Set2 Set1, or CDAB for a final answer in this example). In addition, another symbol was used for the negation of <-//- or -//->.

The order of the three symbols and the ordering operations in each statement were varied across trials. Statements were only presented once, and participants paced themselves through the trials with a limit of 20 seconds per presentation. After all three symbols were presented, participants selected their responses from an alphabetized list (Was & Woltz, 2007). There were 32 sets of stimuli, and the average correct response for each individual was determined. Internal consistency reliability was determined, along with intercorrelations with the other individual measures.

Background Knowledge Measure

This study incorporated background knowledge questions embedded within the reading comprehension assessment. These questions were not specific to the text presented, but they were taken from the same source; thus, discourse topic, type, and

style were matched. Participants answered a total of 18 background knowledge questions from three domains (business, history, and science). Similar measures have been used in both reading and listening comprehension assessments (Hannon & Daneman; 2006; Was, 2005; Was & Woltz, 2007). Was (2005) used these materials with listening comprehension and found internal coefficients ranging from .56 to .63. In the study conducted by Was, true/false statement question format was used; however, the current study converted the literal, inference and background knowledge questions to a short answer format. It should be noted that background knowledge questions were scored along with the target questions.

Vocabulary Ability Measure

This study used the Nelson-Denny Reading Test Form G "Vocabulary" section (Brown, Fishco, & Hanna, 1993). This paper/pencil test contains 80 items. The paper/pencil version is a timed test (17 minutes). In the present study, all items were presented to all participants, regardless of time required to complete them, and total time was recorded. The words are taken from a scaled list. The authors reported that these words were "drawn from current and widely used high school and college texts: to ensure maximum relevance, focus was on words that must be known by students in order to cope successfully with school assignments" (p. 2). These words are used in general writing but are not high frequency words. Beck, McKoewn, and Kucan (2005) described these as Tier 2 words, or those that children use in academic settings and generally cross domains (e.g., coincidence, industrious, absurd). The test manual reported a reliability rating of the vocabulary section as r = .86 with extended study time (Brown, Fishco, & Hanna, 1993).

In the present study, participants were presented with a random order of vocabulary items such as: "Onerous means:" 1. likable, 2. burdensome, 3. flexible, 4. fragile, and 5. matured. (Note, this is not an item used in the measure.). After responding with a key press, participants continued through all of the items. Participants responded to each item, but could not return to previous ones. Scores were determined by percentage correct. Internal consistency reliability was determined, along with intercorrelations with the other individual difference measures. These findings are reported in the Results section.

Reading Comprehension Measure

The reading comprehension measure was designed to be very similar to reading tests participants have taken in the past. Participants read a series of paragraphs extracted from larger texts (i.e., college textbooks), and answered question prompts. The topics of the paragraphs were from the domains of science, history, and business. These topics were chosen because they should not lend themselves to high domain knowledge.

There were six passages from each domain for a total of 18 passages, and each passage had an average of 222 words with a range from 150 to 260 ($\underline{SD} = 30$) words. After each passage, participants responded to short answer questions (2 literal, 2 inference, and 1 background knowledge). Each question required a literal interpretation of the text, or an inference or chain of inferences that linked pieces of information from the text with general world knowledge. The complete reading comprehension assessment is provided in the Appendix.

Participants used a standard keyboard to type in their responses. Each passage was presented on the computer screen in its entirety. Participants were instructed to read

the passage for understanding and told that they would be assessed on their reading comprehension. They answered the questions by typing in their response, then hit "enter" to continue on to the next question or passage. For half of the passages, the text was available while participants answered the questions. The series of passages and questions were presented in random blocks of three passages each.

Background knowledge (BK), literal (L), and inference (I) questions were scored for each of the participants on all 18 passages. Each response was scored as (2) sufficient (1) partial credit, or (0) insufficient.   Each question had a list of acceptable answers determined from the pilot study. One grader scored all responses. An additional grader scored the responses of 30% of the total passages (approximately 4,500 questions). Inter-rater reliability was determined for overall reliability. The passages were also analyzed for item reliability. The results of these analyses are reported in the Results section.

A pilot study was performed to norm the reading comprehension materials and determine their appropriateness. Although the reading comprehension materials have been used in previous studies, the studies were in listening comprehension and used true/false question format (Was, 2005; Was & Woltz, 2007). Forty-seven participants were involved in the pilot study taken from the same subject pool used in the experiment. No participants involved in the pilot study were used in the experiment. The pilot study reading comprehension assessment consisted of 30 passages. However, there were signs of fatigue and the experiment was taking well over 2 hours. Therefore, the number of passages was reduced to 18. The results of the pilot study are reported in the Results section.

Procedures

After being assigned to a carrel, participants signed a consent form that described the study in general terms. As noted earlier, E-Prime software controlled the presentation of the stimulus; however, participants controlled the presentation of the questions. For example, after they read the text they hit the spacebar to have the first question presented. After the question was presented, the participants entered their response with the keyboard and pressed the "enter" key to continue. In half of the trials, the text remained on the screen during the question presentation.

The measures were presented in a counterbalanced block design. The order of individual difference measures (WM, VA, and EB), and the reading measure changed between the twelve versions of the study. For example, in the 1st version WM, was followed by VA, then EB. Then following a short break one version of the reading assessment was given. In version 2, VA was followed by WM then EB, then a second version of the reading assessment. Participants were randomly assigned to one of the 12 versions. The individual difference variable measures were always completed during the 1st hour, followed by a brief break (5 minutes). Participants then performed the reading comprehension measure. The 12 versions were developed to control for order effects. The entire session lasted approximately 2 hours.

Time Elements

As noted, some of the individual measures have time elements involved, but most are self-paced. For the reading comprehension measure, "time measures" were collected from the moment participants were presented with a new passage to the moment they requested the first question. This allowed a test for adherence to the experiment protocol

(Are participants really reading the passage?). Time measures were also collected for each question response to determine if there was a trend for longer reading times as a function of text availability and question type. Time element results are reported in the Results section.

Analyses

This experiment was designed to test the hypotheses that epistemic beliefs (EB), text availability (TA), and question type (QT) interact when readers answer questions on a reading comprehension assessment. It is well established that there are several independent variables related to accuracy on reading comprehension tests. An assumption of a method that uses more than one predictor variable is that the variables are independent of one another (Cohen, 1988). Analyses of correlations of independent measures and the reading assessment were performed to test this assumption for the present study. In addition, analyses of the two epistemic beliefs scales were performed to determine their appropriateness for use in this study. A repeated measures Analysis of Covariance (ANCOVA) was then performed, using levels of the dependent measure with TA, and QT, with EB as independent variables, after entering BK, VA, and WM as covariates. Significance level for all tests was set at .05, unless otherwise noted.

CHAPTER III

RESULTS

Analytic Structure

Separating out epistemic beliefs' predictability of reading comprehension from

other individual difference variables such as background knowledge, vocabulary ability,

and working memory involved four stages of analyses. The first set of analyses was

conducted to determine the reliability of the reading assessment across participants'

reading ability levels, as well as across each domain, question type, and text availability

condition. The second set of analyses consisted of Principal Component Analyses (PCA)

of the two epistemic beliefs scales administered to participants to present evidence of

validity of the scales used in this experiment. The third set of analyses was conducted to

determine not only the reliability of the individual difference measures (i.e., vocabulary

ability, working memory capacity, background knowledge, and epistemic beliefs), but

how they correlated with each other and with the reading assessment. The final set of

analyses consisted of a repeated measures analysis of covariance (ANCOVA) conducted

to determine the relationships between individual difference measures and the reading

assessment. This last set of analyses focused primarily on the nature of the relationship

between individual difference measures considered as covariates (i.e., vocabulary ability,

working memory, and background knowledge) and epistemic beliefs, and their interactions with the independent variables of text availability and question type.

<u>Reliability of Reading Comprehension Assessment</u>

The first set of analyses was conducted to determine the reliability of the reading assessment. Table 2 presents the means and standard deviations for the assessment, overall and as a function of question type and text availability. As noted previously, there were 60 participants in a pilot study who completed a reading comprehension measure with 30 passages. Twelve participants were excluded from the pilot study analysis due to computer problems which created incomplete data. Participants were demonstrating fatigue; therefore the number of passages was reduced to 18 to ensure completion of the entire reading assessment within 1 hour.

For the experiment, analysis was completed on a total of 155 participants. The mean accuracy for the overall assessment was approximately 60%. As expected, literal question accuracy rates were higher when the text was available than when it was not. This was also true for inference questions. Background knowledge question accuracy was the lowest (i.e., approximately 25%). For text availability, the difference between with and without text available conditions was significant, $\underline{t}(154) = 12.603$, $\underline{p} < .000$, Cohen's $\underline{d} = .88$. For question type, there was also a significant difference between the means of literal and inferential questions, $\underline{t}(154) = 18.207$, $\underline{p} < .000$, Cohen's $\underline{d} = .1.23$.

Table 3 presents the means and standard deviations for each of the three domains across question type and text availability conditions. In analyzing the data for the different domains, business domain averages were highest. Business inference questions

Table 2.
Mean Accuracy and Standard Deviations, on Reading Comprehension Assessment

Component		Overall		With Text		Without Text	
	N	Mean	SD	Mean	SD	Mean	SD
Overall Assessment	155	.59	.06				
Literal	155			.80	.14	.67	.18
Inference	155			.62	.15	.55	.17
Background Knowledge	155	.25	.13				

differed slightly as a function of text availability, and this difference was only marginally

significant p = .054. However, business literal questions did differ significantly as a

function of text availability; $t(154)$ = 10.978, p < .000. Participants were more accurate

when the text was made available during question answering. History domain averages

were lowest in background knowledge. History inference questions varied somewhat as a

function of text availability, and this difference was significant, $t(154)$ = 2.77, p = 006.

Participants were more accurate on inference questions when the text was available.

History literal questions showed a similar pattern to business literal questions, in that the

difference as a function of text availability was significant $t(154)$ = 10.015, p < .000.

Participants were more accurate on literal questions when the text was available. The

science domain had similar trends, with a low background knowledge mean. Inference

questions differed slightly more than business as a function of text availability, and this

difference was also significant $t(154)$ = 5.83, p < .000. Participants were more accurate

on science inference questions when the text was available. Science literal questions also

Table 3.
Means and Standard Deviations, on Reading Comprehension Assessment for Three
Domains

Domain-Question Type		Overall		With Text		Without Text	
	N	Mean	SD	Mean	SD	Mean	SD
Business-BK	155			.44	.28	.44	.29
Business-Inference	155			.68	.22	.64	.24
Business-Literal	155			.86	.17	.67	.23
Business-Total	155	.62	.03				
History-BK	155			.10	.18	.11	.21
History-Inference	155			.54	.18	.49	.20
History-Literal	155			.84	.18	.66	.21
History-Total	155	.46	.09				
Science-BK	155			.16	.24	.22	.27
Science-Inference	155			.66	.22	.53	.23
Science-Literal	155			.80	.16	.67	.23
Science-Total	155	.51	.06				

showed a significant difference as a function of text availability; $t(154) = 6.925$, $p <$ .000. Participants were more accurate when the text was made available. This analysis illustrated that participants performed similarly across all three domains. As mentioned previously, there were significant differences between the types of questions asked in that participants were more accurate on literal questions than inference questions in every domain.

The analyses indicated that the reading assessment as a whole was moderately reliable. The Cronbach's alpha for the overall reading assessment was .86. Cronbach's alphas on the measures for background knowledge, inference, and literal questions were .70, .82, and .83, respectively. For business, history, and science domains, the Cronbach's alphas were .91, .82, and .75, respectively. One item in science (passage 9, literal question 2, in the "with" text condition-S9L2W) was excluded from the analysis of the experiment data because it had no variability; all participants responded accurately to the question.

As noted previously, 30% of the data were scored by another individual and compared with the scores given by the researcher. The single measure intraclass correlation for the measure was .97 well within the accepted levels of assessments using short answer protocols (Garnham, 1981).

The analyses indicated that the reading assessment was generally internally consistent. The results indicated that question type differed as expected; literal questions were easier to answer than inferential questions. The results for text availability were somewhat as predicted. Accuracy rates when the text was available were higher than when it was not, and response patterns were consistent across all three domains. There were 12 versions of the reading assessment used in this study, but there were no significant differences in response patterns across different versions, suggesting that there were no order effects.

<u>Factor Analyses of Epistemic Beliefs Scales</u>

A Principal Components Analysis (PCA) was performed on Schommer's (1990) EQ scale. Initial analysis on the EQ scale listed over 23 factors with Eigenvalues greater

than 1. As stated previously, the "Learning is Simple" and "Truth is Certain" components have been used by most reading researchers investigating the effects of epistemic beliefs on reading. The items that load on these two factors are consistent with how individuals learn, and how participants may perceive certain aspects of the nature of truth. Schommer (1990) suggested creating 12 subsets of the scale, and delineating them with four Factor names (Learning is Simple, Learning is Quick, Ability is Fixed, and Truth is Certain). These four factors are then used to create component scores.

Using the two component scores, simple and certain, previous researchers generally coded participants as either naïve or mature in their epistemic beliefs based on their responses to these items of the scale by totaling the scores and determining a median split among participants (Burton & Daneman, 2007; Kardash & Scholes, 1996; Mason, Gava, & Boldrin, 2008; Rukavina & Daneman, 1996). Analysis of data from the current study yielded very low internal consistency ratings (i.e., Cronbach's alphas were less than .30 for the subscales) using this approach. Although factor loadings for some of the items were similar to Schommer's (1990) findings, most were not. It should also be noted that the subscales did not load consistently with Schommer's four factors. Therefore, the creation of the subscales was not used in the current study; an alternative approach for determining component scores for epistemic beliefs was used.

Using all 63 items of Schommer's EQ scale, an Exploratory Factor Analysis was conducted using both pilot and experiment participants ($\underline{N}$=202). Table 4 presents items, factor loadings, communalities, Eigenvalues, and percentages of variance explained for the EQ scale.

Table 4.

EQ Principal Components Analysis Varimax Rotation (*N*=202): Items, Factor Loadings, Communalities, Eigenvalues, and Percentages of Variance

	Item	Factor Loading				Communality
		1	2	3	4	
50	Working hard on a difficult problem for an extended period of time only pays off for really smart students.	**.57**				.35
1	If you are going to be able to understand something, it will make sense to you the first time you hear it.	**.47**				.38
56	A tidy mind is an empty mind.(-)	**-.46**				.24
47	Some people are born good learners; others are stuck with limited ability.	**.46**				.24
51	If a person tries too hard to understand a problem, he or she will most likely just end up being confused.	**.46**				.23
31	Being a good student generally involves memorizing facts.	**.45**				.31
55	Students who are average in school will remain average for the rest of their lives.	.42		.32		.28
63	You will just get confused if you try to integrate new ideas in a textbook with knowledge you already have about a topic.	.41		.39		.34
37	Learning definitions word for word is often necessary to do well on tests.	.40	-.36			.30
48	Nothing is certain but death and taxes. (-)	-.40				.18
39	If a person cannot understand something in a short time, he or she should keep trying. (-)	.40	.30			.26
59	The best thing about science courses is that most problems have only one right answer.	**.39**				.21
19	Educators should know by now which is the best method, lectures or small group discussions.	**.36**				.14
7	I often wonder how much my teachers really know. (-)	-.35				.14
16	Things are simpler than most professors would have you believe.	**.35**				.17

Table 4. Continued

	Item	Factor Loading				Communality
		1	2	3	4	
6	You can believe almost everything you read.	**.33**				.13
3	For success in school, it is best not to ask too many questions.	.33		.30		.22
49	The really smart students do not have to work hard to do well in school.	**.32**				.12
28	Everyone needs to learn how to learn. (-)		**.52**			.27
4	A course in study skills would probably be valuable. (-)		**.50**			.30
60	Learning is a slow process of building knowledge. (-)		**.48**			.26
26	Genius is 10% ability and 90% hard work. (-)		**.45**			.28
43	Getting ahead takes a lot of work. (-)		**.45**			.23
25	Students have a lot of control over how much they can get out of a textbook. (-)		**.41**			.17
22	You never know what a book means unless you know the intent of the author. (-)		**.41**			.24
11	A good teacher's job is to keep his or her students from wandering off the right track.		-.39		.31	.30
32	Wisdom is not knowing the answers, but knowing how to find the answers. (-)		**.39**			.24
30	A sentence has little meaning unless you know the situation in which it is spoken. (-)		**.38**			.20
38	When I study, I look for specific facts.		-.37			.23
15	The most successful people have discovered how to improve their ability to learn. (-)		.36	.31	-.33	.35
24	If I find the lime to reread a textbook chapter, I get a lot more out of it the second time. (-)		**.35**			.18
27	I find it refreshing to think about issues that authorities cannot agree on. (-)			**.55**		.33
13	People who challenge authority are overconfident.	.35		.50		.38

Table 4. Continued

Item		Factor Loading				Communality
		1	2	3	4	
41	If professors would stick to the facts and theorize less, one could get more out of college.	.36		.47		.40
34	Truth is unchanging.			**.47**		.32
23	The most important part of scientific work is original thinking. (-)			**.46**		.25
2	The only thing that is certain is uncertainty itself. (-)			**.45**		.23
35	If a person forgot details but was able to come up with new ideas from a text, I would think they were bright.(-)			**.37**		.17
20	Going over a difficult textbook chapter usually will not help you understand it.			**.37**		.28
46	Often, even advice from experts should be questioned. (-)			**.37**		.23
45	You should evaluate the accuracy of information in a textbook if you are familiar with the topic. (-)		.30	.36		.22
61	Today's facts may be tomorrow's fiction. (-)			**.34**		.16
44	It is a waste of time to work on problems that have no possibility of coming out with a clear-cut and unambiguous answer.			**.33**		.26
12	If scientists try hard enough, they can find the truth about almost anything.				**.52**	.39
8	The ability to learn is innate.				**.51**	.33
53	Usually you can figure out difficult concepts if you eliminate all outside distractions and really concentrate. (-)				**-.50**	.30
58	I appreciate instructors who organize their lectures meticulously and then stick to their plan.				**.48**	.27
21	Scientists can ultimately get to the truth.			-.38	.46	.37
40	Sometimes you have to accept teachers' answers although you do not understand them.				**.43**	.25
36	Whenever I encounter a difficult problem in life, I consult my parents.				**.43**	.19

Table 4. Continued

Item		Factor Loading				Communality
		1	2	3	4	
10	Successful students understand things quickly.	.41			.42	.35
57	An expert is someone who has a special gift in some area.				**.33**	.17
5	How much a person gets out of school mostly depends on the quality of the teacher.				**.32**	.15
33	Most words have one clear meaning.				**.31**	.19
Eigenvalues		4.5	3.6	3.4	3.3	
% of Variance		7.1%	5.7%	5.4%	5.2%	

An observation of the scree plot supported a four- or five-factor solution. Only one item loaded on the fifth factor exclusively; therefore, an analysis set at a four-factor solution with varimax rotation explaining 23% of the variance was used for this study.

Items that yielded factor loadings of .30 or greater were used to construct component scores for each of the four factors. This cutoff point was also used by previous researchers analyzing epistemic belief scales and is an acceptable standard (Schommer, 1990; Schraw et al., 2002; Tabachnick & Fidell, 2007). A similar scoring system was also used by Schraw et al. (2002) in their comparative study of the EQ and the EBI. Constructing scores in this way suggests a unique factor structure for this study's population.

The mean for the items loading on Factor 1 "Ability is Innate" was 2.19,($\underline{SD}$ = .53), using 11 items. The mean for Factor 2 "Learning is Simple" was 1.91, ($\underline{SD}$ = .46), using 10 items. The mean for Factor 3 "Truth is Certain" was 2.66, ($\underline{SD}$ = .53), using 9 items. The mean for Factor 4 "Trust Authority" was 2.69, ($\underline{SD}$ = .52), using 8 items. Note, these scores were based on a 5-point Likert scale. Generally, lower scores indicate more mature epistemic beliefs. The items marked with a (-) following the statement are reverse scored. The reliability coefficient for the overall EQ scale was .69 for 202 participants.

This section provides a discussion of each of the factors in order to demonstrate why the Factor 2 score was chosen as the component to test this study's hypothesis. As indicated in Table 4, for Factor 1 or "Ability is Innate" respondents agreed to statements that learning happens fast or not at all (i.e., *Working hard on a difficult problem for an extended period of time only pays off for really smart students*.). This would suggest that

participants believe people have an innate ability to learn, and hard work and study may be a waste of time for some. Using this component to predict reading comprehension may depend on the participant's perception of their own abilities to learn.

In contrast, Factor 2 or "Learning is Simple," contained items addressing the complexity of learning (i.e., *Learning is a slow process of building knowledge*.). Respondents endorsed statements that supported integrating ideas from across text or classes (i.e., *I try my best to combine information across chapters, or even across classes.*). In addition, learning was viewed as something the learner had control over, and that knowledge could be improved with hard work or study skills classes. The score from this factor was used to analyze the effect of epistemic beliefs on reading comprehension in this study because of its inherent tie to how participants view learning. This will be discussed further in the following section.

For items loading under Factor 3, "Knowledge is Certain," participants may read texts at superficial levels looking for the facts presented in the reading (i.e., *You will just get confused if you try to integrate new ideas in a textbook with knowledge you already have about a topic.*). Contrary to this, Factor 3 also had several items that refer to knowledge and truth as being 'unchanging' or ambiguous (i.e., *The only thing certain is uncertainty itself.*). Further, other items that loaded on this factor find readers will delve more deeply into what the text is trying to convey (i.e., *I find it refreshing to think about issues that authorities cannot agree on.*). This seemingly contradictory approach to knowledge and truth renders the factor difficult to interpret with regard to reading without a follow up interview of how participants thought through the selected statements.

Finally, for items loading on Factor 4, "Trust Authority," participants supported the notion that scientists and teachers can or should be trusted (i.e., *Sometimes you have to accept teachers' answers although you do not understand them.)*. This component is difficult to interpret in terms of how a person would be engaged with the reading process. Factor 4 is also rendered difficult to interpret given the high loading of a statement about innate ability (i.e., *The ability to learn is innate.)*. This statement would logically fall under Factor 1; therefore, it is difficult to determine why it loaded on this factor. Additionally, participants endorsed avoiding integration, and working too hard on problems. It could be interpreted that readers are going to avoid integration and only process on a text-based level, more often than trying to integrate information across text. This conjecture is supported by a negative factor loading on an item referencing hard work (i.e., *Usually you can figure out difficult concepts if you eliminate all outside distractions and really concentrate.)*.

In analyzing the four factors from the EQ scale, the most interpretable factor is Factor 2 "Learning is Simple." The statements endorsed a mature treatment of learning and knowledge. This factor was therefore selected to test the current study's hypothesis that participants with more mature epistemic beliefs engage in deeper reading of expository text. It may be that these participants have developed reading strategies that allow them to develop better situation model representations of text than participants with more naïve epistemic beliefs.

As noted previously, a second epistemic belief scale, the Epistemic Beliefs Inventory (EBI) developed by Schraw et al. (2002), was added to the study to validate the findings of the EQ scale. This scale contains 7 items taken from the EQ scale and an

additional 21 separate items for a total of 28 items, and it was administered to 146 participants. It uses the same Likert scale (i.e., 1 to 5), and lower scores indicate more mature epistemic beliefs, The EBI scale was analyzed in a similar fashion to the EQ scale. A Principal Components Analysis (PCA) was conducted, and initially the data yielded 10 factors with Eigenvalues over 1, which explained 63% of the variance. An examination of the scree plot suggested a four- or five-factor solution. A follow-up PCA was conducted with five factors that explained 40% of the variance; however, Factor 5 had only one item that loaded exclusively on it. A final PCA with varimax rotation was conducted with four factors extracted, explaining 35% of the variance. Table 5 presents the items, factor loadings, communalities, Eigenvalues, and percentages of variance for the scale. As with the EQ scale, the current study created component scores for each factor by taking the average of items that loaded on the specific factors. The scores are comprised of items indicated in bold in Table 5. Items that loaded on more than one factor were excluded from the scores.

The names given to each factor were taken from the Schraw et al. (2002) study, and they are consistent with the items loading under these factors in their study. The mean for Factor 1 "Innate Ability" was 2.86, (SD = .55), using 7 items. The mean for Factor 2 "Simple Knowledge" was 2.2, (SD = .47), using 6 items. The mean for Factor 3 "Omniscient Authority" was 2.9, (SD = .52), using 6 items. The mean for Factor 4 "Certain Knowledge" was 3.1, (SD = .75), using 4 items. The items marked with a (-) following the statement are reversed scored. The reliability coefficient for the overall EBI scale was .63 for 146 participants.

Table 5. EBI Principal Components Analysis Varimax Rotation (*N*=146): Items, Factor Loadings, Communalities, Eigenvalues, and Percentages of Variance

	Item	Factor Loading				Communality
		1	2	3	4	
75	Some people just have a knack for learning.	.64				.42
80	Smart people are born that way.	.55				.31
10	Successful students understand things quickly.	.52		.32		.44
70	Too many theories just complicate things.	.51				.33
49	The really smart students do not have to work too hard to do well in school.	.48				.33
73	How well you do in school depends on how smart you are.	.48				.35
72	Some people are born with special gifts and talents	.46	-.42			.46
67	Some people will never be smart no matter how hard they work.	.46	.31			.34
69	Parents should teach their children all there is to know about life.	.33				.14
20	Going over a difficult textbook chapter usually will not help you understand it.		.62			.43
83	Working on a problem with no quick solution is a waste of time.		.61			.43
64	Most things worth knowing are easy to understand.		.56			.38
74	If you don't learn something quickly, you won't ever learn it.		.50			.27
78	The more you know about a topic, the more there is to know.		-.50			.29
41	If professors would stick to the facts and theorize less, one could get more out of college.	.45	.46			.43
76	If two people are arguing about something, at least one of them must be wrong.		.41			.27
66	People should always obey the law.			.76		.60
82	People who question authority are trouble makers.			.65		.48
81	When someone in authority tells me what to do, I usually do it.			.62		.39
77	Children should be allowed to question their parents' authority.			-.43		.28

Table 5. Continued

Item	Factor Loading				Communality
	1	2	3	4	
79 What is true today will be true tomorrow.			**.42**		.27
59 The best thing about science courses is that most problems have only one right answer.			**.31**		.15
84 Sometimes there are no right answers to life's big problems.				**.70**	.51
68 Absolute moral truth does not exist.			-.34	**.60**	.48
65 What is true is a matter of opinion.			-.32	**.54**	.41
51 If a person tries too hard to understand a problem, he or she will most likely just end up being confused.	.30	.35		.54	.50
16 Things are simpler than most professors would have you believe.				**.33**	.15
Eigenvalues	3.63	2.33	2.07	1.86	
Percentage of Variance	10%	9.5%	8.3%	7.5%	

Unfortunately, the statements from the EBI scale were more difficult to interpret than those on the EQ scale, particularly with regard to learning and the nature of knowledge. Additionally, the authors used seven of Schommer's EQ items. As a result of these issues, data from this scale were not used in the study to test the hypotheses. However, the findings do provide evidence that the EQ and EBI scales are somewhat parallel in measuring the construct of epistemic beliefs. The correlation between the two scales was moderate but significant, $\underline{r}$ = .55, $\underline{p}$ < .01. As Schraw et al. (2002) noted in their study, further refinement is required of these scales. It is interesting to note that in the Schraw et al. study, their analyses of the factors yielded moderate correlations to a reading comprehension measure ($\underline{r}$ ranging from = .36 to -.29). In the current study, correlations with reading comprehension scores and factors of the EBI scale were low and negative ($\underline{r}$ ranging from -.04 to -.23). It should be noted that Schraw et al. used a multiple choice inference only protocol in the reading comprehension assessment in their study.

<div align="center">Reliability of Individual Measures</div>

The purpose of the next set of analyses was to determine reliability of the individual measures (i.e., vocabulary ability, working memory capacity, background knowledge, and epistemic beliefs), and correlations with the reading assessment and subsets within the measures. Table 6 presents the means, standard deviations, and reliability coefficients for the individual measures. The vocabulary ability assessment mean and reliability was consistent with the information reported by the authors of the Nelson-Denny Vocabulary Assessment for college students (i.e., mean = .80, and α = .77; Brown, Fishco, & Hanna, 1993). The working memory measure mean and reliability

Table 6.
Means, Standard Deviations, and Reliability Coefficients for Individual Measures

Measure	N	Mean	SD	Cronbach's α
Vocabulary	202*	.802	*.14*	.84
Working Memory	202*	.655	*.21*	.62
EQ-Learn/Simple	202*	1.91	*.46*	.69
EQ-Trust Authority	202*	2.69	*.52*	.55
EQ-Truth is Certain	202*	2.27	*.53*	.62
EQ-Ability is Innate	202*	2.19	*.53*	.59
EQ-Knowledge is Simple	202*	2.66	*.55*	.60
EQ-Total	202*	2.55	*.36*	.70
EBI-Ability is Innate	146**	2.40	*.60*	.57
EBI-Omniscient Authority	146**	2.92	*.57*	.60
EBI-Learning is Quick	146**	2.19	*.44*	.40
EBI-Knowledge is Certain	146**	2.81	*.43*	.51
EBI-Total	146**	2.59	*.55*	.70

*Analysis completed using pilot participants.
**Scale added after preliminary analysis.

coefficient was moderate, which was also consistent with other measures of this type. For example, Kyllomen and Christal (1990) reported a Cronbach's α of .65 on a similar working memory task. Reliability estimates for the epistemic belief scales and subscales were moderate, but consistent with coefficients provided in previous studies (Schommer, 1990; Schraw et al. 2002). The first three variables listed in Table 6 were used in the current study.

Information regarding the other factors for the epistemic beliefs scales was provided to support the concurrent validity of the EBI and EQ scales, but was not used in subsequent analyses. As noted earlier, there were three individual differences variables that needed to be accounted for in this study (i.e., background knowledge, vocabulary ability, and working memory), prior to analyzing the variable of interest (epistemic beliefs). Table 7 presents the intercorrelations (Pearson r) among these measures and the reading assessment components. As noted previously, there are five components of the reading assessment: background knowledge (BK); inference questions with text available (WI) and without text available (WOI); and literal questions with text available (WL) and without text available (WOL). Vocabulary ability (VA) had positive, moderate correlations with the reading assessment components (r ranging from .28 to .36), and

Table 7.
Correlations for Individual Measures

Measure	1	2	3	4	5	6	7	8
1 VA	(.84)							
2 WM	.28**	(.62)						
3 BK	.32**	.13	(.70)					
4 WI	.36**	.09	.56**	(.54)				
5 WL	.28**	.10	.33**	.52**	(.52)			
6 WOI	.34**	.18*	.60**	.58**	.48**	(.65)		
7 WOL	.36**	.17*	.51**	.62**	.56**	.69**	(.67)	
8 EQ-SimpLearn	-.19*	-.12	-.16*	-.08	-.14	-.08	-.22**	(.69)

*Correlation is significant at the .05 level 2-tailed.
**Correlation is significant at the .01 level 2-tailed.

these correlations were significant. Working memory (WM) was also positively correlated with the reading assessment components, but the correlations were only significant in the without text conditions. Vocabulary ability and working memory were significant and moderately positively correlated with each other. Other researchers using this type of working memory measure (Was & Woltz, 2007) have reported similar correlations (e.g., vocabulary, $\underline{r} = .21$, listening comprehension, $\underline{r} = .35$). Background knowledge (BK) was significantly positively correlated with each component of the reading assessment at moderate levels ($\underline{r}$ ranging from .33 to .60). This is not surprising, given that background knowledge scores came from the same measurement, using the same type of protocol. It was also anticipated that readers with high background knowledge in a specific domain should score higher in that domain on the reading assessment. BK had a significant moderate, positive correlation with vocabulary ability, but BK and working memory were not significantly correlated. In addition, BK had a small but significant, negative correlation with the component of the epistemic beliefs scale of interest in this study. Correlations between EQ-Simplearn and vocabulary ability and background knowledge were small and negative, but significant. The correlation between EQ-Simplearn and working memory was not significant. The correlations between the reading assessment components and EQ-Simplearn were lower than expected; $\underline{r}$ ranged from -.08 to -.22. The only significant correlation was between EQ-Simplearn and the without literal (WOL) condition. It was anticipated that higher negative correlations would be obtained in the "without" text conditions. Readers with more mature epistemic beliefs (lower EQ scores) would be able to engage in a deeper meaning construction of the text than naïve readers, and this would be more evident in

the "without" text condition where reliance on rereading when answering the questions was not possible. Difficulty in answering inference questions did not tend to differ whether or not the text was available. It appears that there is a relationship between epistemic beliefs and reading comprehension; however, this relationship may be more complex than was predicted.

<u>Repeated Measures Analysis of Covariance</u>

To analyze the effect of epistemic beliefs on reading comprehension, a repeated measures analysis of covariance ANCOVA was performed on the data. Three individual differences measures were included as covariates (vocabulary ability, working memory, and background knowledge), and epistemic beliefs (EQ-Simplearn), text availability (TA), and question type (QT) were entered as independent variables. Table 8 presents the results of the repeated measures ANCOVA. There was an equal sample size for all variables; only participants who completed the final version of the reading assessment were included ($\underline{N}$ = 155). As noted previously, there were 12 versions of the reading assessment created for counter- balancing that varied the order of passages and conditions. Thus, version was included as a between subjects variable to test for any effect of order. For both text availability and question type, there were no interactions with version as expected. This suggests that there were no effects for order.

The ANCOVA revealed a significant main effect for text availability, $\underline{F}$ (1,138) = 26.37, $\underline{p}$ = .000, η_p^2= .16. Participants were more accurate when the text was available $\underline{M}$ = .75, $\underline{SD}$ = .12, than when it was not $\underline{M}$ = .61, $\underline{SD}$ = .15. This could indicate that participants were using a strategy of rereading portions of the text pertaining to the questions.

Table 8.

Repeated Measures Analysis of Covariance of Epistemic Beliefs (EB) as a Function of Text Availability (TA) and Question Type (QT) with VA, WM, and BK as Covariates

Source	Type III Sum of Squares	Df	Mean Square	F	Partial Eta Squared
TA	.32	1	.32	26.37***	.16
VA*TA	.12	1	.12	9.83**	.07
WM*TA	.01	1	.01	.95	.01
BK*TA	.04	1	.04	3.12	.02
EB*TA	.00	1	.00	.01	.00
Version*TA	.16	11	.02	1.20	.09
Error(TA)	1.67	138	.01		
QT	.45	1	.43	46.42***	.25
VA*QT	.05	1	.05	5.07**	.04
WM*QT	.01	1	.01	.76	.01
BK*QT	.15	1	.15	16.68***	.11
EB*QT	.10	1	.10	10.85***	.07
Version*QT	.12	11	.01	1.18	.09
Error(QT)	1.27	138	.01		
TA*QT	.01	1	.01	.98	.01
TA*QT*EB	.04	1	.04	2.48	.02

VA = Vocabulary Ability; WM = Working Memory; BK = Background Knowledge.
**$p < .05$, ***$p < .001$

The interaction of text availability with vocabulary ability was significant, F (1,138) = 9.83, p = .002, η_p^2= .07. Vocabulary had a significant influence on participants' ability to answer questions. This indicates that participants were able to answer reading comprehension questions more accurately based on their overall proficiency in vocabulary, particularly when the text was not available. This was the only individual difference variable that interacted with text availability; however, the interaction of background knowledge with text availability did approach significant levels, p = .08. Participants with higher background knowledge scores were marginally more accurate than those without background knowledge when the text was not available. Text availability did not interact with working memory and epistemic beliefs.

Additionally, the ANCOVA revealed a significant main effect for question type $\underline{F}(1,138) = 46.42$, $\underline{p} = .000$, $\eta_p^2 = .252$. Participants were more accurate on literal questions $\underline{M} = .75$, $\underline{SD} = .13$, than inference questions $\underline{M} = .59$, $\underline{SD} = .13$. There was a significant interaction between question type and background knowledge, $\underline{F}(1,138) = 16.675$, $\underline{p} = .000$, $\eta_p^2 = .11$. When participants had background knowledge in a particular subject, they were able to respond more accurately to literal questions than those who had little background knowledge. In addition, they were able to respond more accurately to inference questions about that subject as compared to those with little background knowledge. It should also be noted that the relationship between background knowledge and inference questions was stronger than literal. This will be discussed further in the Discussion section.

There was also a significant interaction between question type and vocabulary ability, $\underline{F}(1,138) = 5.07$, $\underline{p} = .026$, $\eta_p^2 = .04$. This suggests that participants' vocabulary ability influenced reading comprehension when comparing literal and inference questions. This relationship is more robust with inference questions than literal questions. Again, question type did not interact with working memory.

The component of the epistemic beliefs scale (EQ-Simplearn) did provide some explanation of differences in questions type, $\underline{F}(1,138) = 10.854$, $\underline{p} = .001$, $\eta_p^2 = .07$. The interaction of epistemic beliefs with question type was as predicted; however, it was anticipated that the largest difference would be in the accuracy on inference questions. Instead, epistemic beliefs had a stronger relationship with literal questions than inferential questions. This suggests that participants' ability to understand text at the literal level is

influenced by their perspectives of learning. Further, it was expected that text availability and question type would interact; however, this did not occur.

This study showed that a person's individual epistemic beliefs about knowledge and learning may provide some means for determining performance on reading assessments. It is unclear, however, if the epistemic beliefs scales have the necessary internal consistency required to make their use viable. The factor analyses of the epistemic beliefs scales provided minimal evidence that readers' perspectives of knowledge and learning influence reading comprehension. Although it may make sense theoretically, statistical evidence is lacking.

With regard to time elements collected for the reading assessment. In a 2 X 2 ANCOVA performed for question response time there was a significant main effect for question type, $\underline{F}(1,137) = 10.437$, $\underline{p} = .002$, $\eta_p^2 = .071$. Average times for question response were about twice as long for inference questions than for literal questions. There was no main effect for text availability or question type. Further, there was no significant interaction between text availability and question type.

CHAPTER IV

DISCUSSION

<u>Relationships Among Measures</u>

The present experiment explored the relationship between epistemic beliefs and test structure variables of reading comprehension assessments using the CI model as a theoretical framework (Kintsch, 1998). Specifically, the study explored the interaction of epistemic beliefs with text availability and question type on a reading comprehension assessment. The CI model suggests that readers need to integrate facts presented in the text with their general background knowledge to form viable situation models that can be instantiated into long term memory. Participants who hold mature epistemic beliefs with regard to the complexity of learning may read at a deeper level than naïve readers, because they might have developed reading strategies for remembering the text better than naïve participants. This strategic reading would be more evident when the text was not available, because mature readers had formed more substantial textbase and situation model representations of the texts than naïve readers. The results of the study suggested, however, that all readers had a difficult time responding to both types of questions when the text was not available. Short answer format requires participants to generate responses rather than select from a set of possible choices as in multiple choice format assessments. The findings from this study do not support the prediction that participants with mature epistemic beliefs would outperform naïve readers when the text was not

available during question answering. There was no interaction between text availability and epistemic beliefs.

Based on the assumptions of Kintsch's (1998) CI model, it was predicted that literal questions would be easier to answer than inference questions. It was also predicted that this effect of question type would differ as a function of readers' epistemic beliefs. Mature readers should be able to apply their beliefs with regard to the complexity of learning when reading expository text from various domains, and they should be able to do this at both surface (textbase) and situation model levels of text comprehension. Thus, the difference between literal and inference questions should be much smaller for mature readers than for naïve readers. The results of the present study did not conform to this prediction. It appeared that mature readers used effective reading strategies to help them answer questions more accurately at the literal level than naïve readers, but this difference was not as evident with inference questions. It may just be that short answer inference questions are difficult to answer irrespective of readers' perspectives on learning, or it may also be that perspectives on learning do not predict strategies used when reading.

In addition to epistemic beliefs, this study also examined the role of several other individual difference variables associated with reading comprehension. Previous studies have suggested that mature epistemic beliefs can mitigate the detrimental effects of low working memory capacity (Burton & Daneman, 2007; Daneman & Hannon, 2001). For example, Burton and Daneman used a reading span task and a digit span working memory task to compare results on the SAT-Verbal between participants who held mature versus naïve epistemic beliefs. They found that participants who held mature

epistemic beliefs and rated high on the reading span task performed the highest on reading comprehension. However, their forward digit span task was not a significant predictor of reading comprehension. The current study used an attention-based measure of WM. Similar to the findings of Daneman and colleagues, the results of this study indicated that if only attention factors are taken into consideration, working memory's ability to predict reading comprehension accuracy drops dramatically. Working memory did not interact with either of the test structure variables (i.e., text availability and question type) included in this study. It should be noted that the working memory measure uses a multiple choice format, unlike the reading comprehension assessment which uses a generated response format. The different formats may help explain why the results were not similar to those of previous studies.

The CI model (Kintsch, 1998) provides a framework for understanding the levels of text representation readers engage in during reading. This study provided evidence that background knowledge influences accuracy on inference questions to a greater degree than literal questions. There was an interaction between background knowledge and question type. Participants with high background knowledge were more accurate than those with low background knowledge on inference questions. This was also true of literal questions without the text, but to a lesser degree. Further, background knowledge also tends to influence reading comprehension when text is not available to a higher degree than when it is available. Readers with high background knowledge were more accurate without the text than readers with low background knowledge; however, this interaction was only marginally significant.

One argument maintained with previous studies is that vocabulary ability is a separate construct from reading comprehension, yet many standardized tests measure them together (Pearson & Hamm, 2005). The findings of this study somewhat support this argument. Vocabulary ability did interact with text availability, but this interaction only explained 10% of the variance. Readers with a higher percentage correct on the vocabulary measure were more accurate on the reading comprehension assessment irrespective of text availability. Vocabulary ability also had a significant interaction with question type, but only 4% of the variance was accounted for by this interaction. Participants with a high vocabulary were able to accurately construct both literal and inference responses, better than those scoring low on the vocabulary measure; however, the interaction was much lower than reported in previous studies. This suggests that there is a relationship between vocabulary ability and reading comprehension, but it may not be as robust when the formats of the measures are not similar.

In this study the main effects of text availability and question type accounted for a small amount of variance on the reading assessment measure. Further, there was no interaction between text availability and question type. The individual difference variables only accounted for a small amount of the variance explained, particularly epistemic beliefs. In the current study, epistemic beliefs was treated as a continuous variable in the analysis, and its effect on readers' abilities to answer literal and inference questions was significant, but minimal. The effect size for the interaction of EB and QT was 7%. With respect to the CI model, this means that there was a small but significant difference between mature versus naïve readers with respect to types of questions asked (i.e., literal versus inference). The interaction was mainly due to literal questions rather

than inference questions, contrary to the predictions of the CI model. One explanation for these results may be that the short answer format used in this study produced lower correlations with individual measures than in previous studies that used multiple choice formats. Further, the use of three domains to measure general reading comprehension indicates that participants may process text from science, business and history differently, thus creating different sources of measurement error.

These findings were somewhat inconsistent with previous studies that examined the predictive power of epistemic beliefs on reading comprehension (Burton & Daneman, 2007; Kardash & Scholes, 1996; Mason, Gava, & Boldrin, 2008; Rukavian & Daneman, 1996; Schommer, 1990). For example, Burton and Daneman reported an effect size of 12% for epistemic beliefs on reading comprehension. One explanation for the difference in effect sizes between the present study and previous work may be that previous research on the role of epistemic beliefs studied this variable in isolation or in combination with only one or two other individual differences or test structure variables. Vocabulary ability, working memory, and background knowledge measured with similar question formats to the reading comprehension measure have been found to be good predictors of reading comprehension. However, in the current study, the small influences of vocabulary ability, background knowledge, and epistemic beliefs, and their interactions with test structure variables indicate that it is important to take into account a number of individual differences and test structure variables to determine their unique relationships with each other.

Another difference between the present study and previous studies on epistemic beliefs and reading may be that most previous studies used predetermined subscales of

Schommer's EQ scale (i.e., "Knowledge is Certain" and "Simple Knowledge") . The current study performed a PCA that yielded very moderate to low factor loadings for these subscales. Further, in comparison to Schommer's study using the EQ scale, the suggested subscales did not load on specific factors and had very low internal consistency. Adding the EBI scale developed by Schraw et al. (2002) provided no significant criterion validity for the relationship to reading comprehension, although the factors did load on similar items. The analysis of the EBI showed low correlations with the other individual difference measures, and a difference in valence with the reading measure. Schraw et al. reported a low positive correlation to the epistemic beliefs factors, whereas this study reported low negative correlations. In addition, the EBI items were hard to interpret with regard to the complexity of learning. For the current study, a composite score of Factor 2 "Learning is Simple" was used to test epistemic beliefs' relationship to reading comprehension. A further explanation of the difference in this study with previous studies may be that mature readers have the ability to discriminate correct answers from reasonable distractors. Based on the current study and previous research, it is difficult to determine whether the EQ and EBI scale actually measure this ability.

<div align="center">Limitations and Future Directions</div>

There are a few limitations of this study. First, only a college age population was used, and people who hold mature epistemic beliefs in the general population may vary a great deal from this sample. Further, the population was predominantly upper class students (e.g., juniors, seniors, and graduates made up 75% of the sample). It may be that this population did not provide the variability needed with the epistemic beliefs scales to

mirror previous studies. Another limitation was the difficulty of the task selected to test reading comprehension. The results suggested that responding to literal questions without the text available was very difficult, and this may be irrespective of a person's epistemic beliefs. Perhaps testing with a multiple choice task and comparing the responses with short answer format would provide a better test of comprehension strategies used when reading.

Given the small amount of variance explained in this study by the individual difference measures on reading comprehension, future studies of reading must take into account several aspects. With regard to accounting for individual differences, background knowledge and vocabulary ability must be taken into consideration when analyzing reading comprehension, because of the unique nature of the cognitive processes that take place during reading. The magnitude of the correlations of the individual measures and the reading assessment were lower than in previous studies indicating that the reading measure used in this study has unexplained measurement error. A study comparing formats may provide further evidence of the unique relationships between test structures and individual measures.

However, with regard to epistemic beliefs, researchers may want to develop a self-report measure that specifically targets strategies used when reading rather than a measure of readers' perceptions of learning from text. In addition, interest in text and motivation were not taken into consideration in this study. Future research should focus on these aspects of reading comprehension. With regard to test structures, the design should provide a multiplicity of domains and topics, and an equal number of literal and inferential questions. Although there was measurement error in the current assessment,

the variability of topics provides a much broader understanding of comprehension in general.

In addition, short answer format proved to be cognitively demanding for the reader, but does it significantly differ from the more traditionally used multiple choice format? A follow-up study comparing these two formats could provide evidence for this argument.

<div align="center">Conclusion</div>

In conclusion, this study's results indicate that a variety of individual differences variables, including epistemic beliefs, and test structure variables interact to influence reading comprehension. However, these effects were smaller than reported in the previous research literature. This may be in part a function of the specific measures used to assess the individual differences variables, or the test structure variables chosen for use in this study. This study may serve as an impetus for future research to investigate the relationships of these types of variables more carefully when assessing reading comprehension, and to investigate other variables that may have more predictive power in reading comprehension assessments.

APPENDIX

READING COMPREHENSION ASSESSMENT

Business (1)
Nearly all decision involve trade-offs; there are advantages and disadvantages, costs and benefits, associated with every action and every choice. A key concept that recurs again and again in analyzing the decision-making process is the notion of opportunity cost. The full cost of making a specific choice includes what we give up by not making the alternative choice. That which we forego, or give up, when we make a choice or a decision is called the opportunity cost of that decision. The concept applies to individuals, businesses, and entire societies. If you decide to take time off in lieu of working, the opportunity cost of your leisure is the pay you would have earned. Part of the cost of a college education is the income you could have earned by working full time instead of going to school. If a firm purchases a new piece of equipment for $3000, there is an opportunity cost; that $3000 could have been deposited in an interest-earning account or lent to another firm. The reason that opportunity costs arise is that resources are scarce. Scarce means limited. Consider one or our most important resources - time. There are only 24 hours in a day, and we must live our lives under this constraint. Many things in life are scarce, and much of economics is considered with behavior in the face of scarcity. If your neighbor mows his lawn today, he won't have time to take his children to the zoo, and this too is an opportunity cost of mowing the lawn.

What is the opportunity cost regarding vacation from work? (Inf.)
Lost wages or vacation pay

What is the effect when resources are scarce? (Lit.)
There is a rise in opportunity cost.

Name three resources involved in opportunity costs? (Inf.)
Time, money, enjoyment

Who is affected by the notion of opportunity costs? (Lit.)
Individuals, businesses, entire economies

How do businesses analyze opportunity costs? (BK)
Opportunity ratio to opportunity property ratio

Business (2)
The least complex and most common form a business can take is the simple proprietorship. There is no legal process involved in starting a proprietorship. You simply start operating. You must however, keep records of revenue and costs and pay personal income taxes on your profit. A professor who does consulting on the side, for example, receives fees and has costs (computer expenses, research materials, and so forth). This consulting business is a proprietorship, even though the proprietor is the only employee and the business is very limited. A large restaurant that employs hundreds of people may also be a proprietorship if it is owned by a single person. Most doctors and lawyers in private practice report their income and expenses as proprietors. In a proprietorship, one person owns the firm. In a sense, that person is the firm. If the firm owes money, the proprietor owes the money; if the firm earns a profit, the proprietor earns a profit. There is no limit to the proprietor's responsibility; if the business gets into financial trouble, the proprietor alone is liable. That is, if a business does poorly and ends up in debt, those debts are the proprietor's personal responsibility. There is no wall of protection between a proprietor and her business, as there are between corporations and their owners.

Regardless of size, what determines a proprietorship? (Inf.)
The number of owners

What is the legal process for starting a proprietorship? (Lit.)
There is no legal process

In a proprietorship, who owns the assets of the firm? (Lit.)
The proprietor, single owner of business

What happens in a bankruptcy, for the proprietor? (Inf.)
The proprietor will be in personal bankruptcy.

What happens to the assets held jointly with a spouse in a bankruptcy of a proprietorship? (BK)
The assets will be considered part of the spouse's assets and be used to relieve debt.

Business (3)
Market penetration means trying to increase sales of a firm's present products in its present markets, probably through a more aggressive marketing mix. The firm may try to increase the customers' rate of use, attract their competitors' customers or current nonusers. For example, Visa increased advertising to encourage customers to use its credit card when they traveled and to switch from using American Express. New promotion appeals alone may not be effective. A firm may need to add more stores, or add short-term price reductions or coupon offers. MCI increased advertising and offered special discounts to encourage consumers to choose MCI over AT&T. AT&T in turn offered MCI customers the chance to 'come back free.' Diversification means moving into totally different lines of business perhaps entirely unfamiliar products. When Japanese based Sony purchased US based CBS records, it expanded from producing electronic equipment into producing music as well, and it is considering other moves that will take it even further from its traditional business. Diversification presents the most challenging opportunities, because it involves both new products and new markets. The further the opportunity from what the firm is already doing the more attractive it may look to the optimists and the harder it will be to evaluate. Opportunities very different from a firm's current experiences involve higher risks. The landscape is littered with failed efforts at diversification. For example, Holiday Corporation learned fast that making mattresses like the ones used in its Holiday Inn motels was NOT one of its strengths.

What is one way that companies can diversify? (Lit.)
They can develop new practices and new markets.

How can a company know if diversifying into a new line will be successful? (Inf.)
Analyze how close a new product is to their existing market.

How would a car dealership diversify? (Inf.)
Opening a service department

How did MCI lure customers away from their competitor? (Lit.)
Offered special discounts

What does diversification usually entail? (BK)
Changing the scope of what the company markets.

Business (4)

The rapid pace of technological change opens up new opportunities, but it also poses challenges for marketers. For many firms success hinges on how quickly new ideas can be brought to market. It's easy for a firm to slip into a production orientation in the flush of excitement that follows a new discovery in a research and development lab. That makes it more important than ever for marketing thinking to guide the production process: starting at the beginning with decisions about where basic research and development effort will be focused. Marketers must also help their firms decide what technical developments are ethically acceptable. For example, many firms have now installed a system to identify the telephone number of an incoming telephone call. It is possible for a firm to know what customer is calling, as well as detailed information about what the customer purchased in the past. This is a very powerful technology, but many people feel this is an invasion of privacy. Similarly, with the growing concern about environmental pollution and the quality of life, some attractive technological developments may be rejected because of their long-term effects on the environment. Aseptic drink boxes, for example, are very convenient but difficult to recycle. In a case like this, what is good for the firm and some customers may not be good for the cultural and social environment or acceptable in the political and legal environments. Being close to the market should give marketers a better feel for the current trends and help firms avoid serious mistakes.

Why is it important for marketing to guide the production process? (Inf.)
The product has to have a market in order to be successful.

What makes up a product's market? (BK)
The set of actual and potential buyers of a product

What is an example of a product that is not good for the environment? (Lit.)
Aseptic drink boxes

What could happen when Research and Development acts before considering all parties? (Inf.)
Failure of the product and losses of revenue

With today's technical advances, what ethical issue must companies address? (Lit.)
Privacy, health hazards

Business (5)

Marketing means different things to practitioners of the discipline. Some believe that it is primarily a matter of promoting products and services to potential customers. Others see it as creating new products and services to serve those customers, or creating innovative and efficient channels of distribution by which to make those products and services available in convenient and price-sensitive ways. Still there are those who live by a principle founded on the hard rock of intelligent pricing - price it right and build a customer base overnight. All of these meanings are relevant to the world of marketing. Most contemporary definitions of marketing seem to fall squarely between the aspects of human nature and market evolution. That is, marketing grows out of our own personal natures and is given an impersonal external life which is propelled by forces that seem beyond our ability to control, or even sometimes to understand. As an example, early proponents of materialism in the late 18th century and early 19th century America agreed with Alexander Hamilton that man's nature is acquisitive. That we are naturally envious, that if we give free reign to our natural impulses, then our own self-interested efforts will serve public good and private ambition. Combined with science and its application to industrial progress, private ambition thus becomes economic entrepreneurialism.

Where does the term marketing stem from? (BK)
It is a term generally referred to as promotion, distribution, and development of products.

Who suggested that man's nature is acquisitive? (Lit.)
Alexander Hamilton

How do materialists perceive private ambition's influence on public good? (Inf.)
Our natural tendencies toward envy and self-interest promote materialism.

What is the effect of self-interest on the business community? (Inf.)
The business community competes for consumers therefore keeping prices down.

Where do contemporary definitions of marketing fall? (Lit.)
Squarely between human nature and market evolution

Business (6)

The foundations of management are, in a large part, built on the dual concepts of efficient organizational structure and effective employee motivation. Andrew Carnegie, whose stewardship of the early steel industry made him one of the richest men in America, was always aware of the part that money plays in the balancing of these concepts. If the firm is not responsive to the need for fair compensation as a critical motivator, then no organization structure is going to remain intact when labor strikes or walks off the job. Carnegie's investigation of the labor question produced a balanced document which affirms the dignity of labor and the purposes of capital or management. It is this relationship between labor and capital that must, according to Carnegie, evolve for the sake of the community. He recognized that good management should produce an environment where strikes and boycotts would be unimaginable. In the absence of such, he sketches a plan for peaceful relations between labor and capital that serves, even today, as a foundation for conduct and arbitrations.

How did Carnegie seek an environment that minimized strikes and boycotts? (Inf.)
Brought management and labor together to see each others' viewpoint

What is the name of groups organized to fight for labor rights? (BK)
Labor unions

What industry was Andrew Carnegie involved with? (Lit.)
Steel

What dual concepts are the foundations to management? (Lit.)
Efficient organizational structure and effective employee motivation

Why should firms be responsive to the need for fair compensation of their labor force? (Inf.)
To minimize the risk of strikes and boycotts that will lose them money

Science (l)
An atom is the smallest particle of an element. An element is a substance made up of only one kind of atom. An atom's mass is made up of three main subparticles called the proton, neutron, and electron. The center of the atom is called the nucleus, and contains the proton and neutron. The nucleus contains most of the mass of the atom, although its volume is much greater. There are usually the same number of protons and neutrons. When there are more neutrons than protons the atom is called an isotope. The proton has a positive charge, the neutron no charge, and the electron has a negative charge. Together, all of the electrons of an atom create a negative charge that balances the positive charge of the protons in the atomic nucleus. Electrons are extremely small compared to all of the other parts of the atom. The mass of an electron is almost 1,000 times smaller than the mass of a proton. Atoms have an atomic number. This is the number of protons in the atom. The atomic weight of the atom is the sum of the proton and neutron in the atom's nucleus. An atom is named by the element it represents. For example, an oxygen atom is represented by the symbol 'O' a hydrogen atom by 'H' and so on.

Where do electrons travel in relationship to the nucleus? (Inf.)
Around the outside of the nucleus

What is an atom called that has more neutrons than protons? (Lit.)
Isotope

What is the chemical symbol for table salt? (BK)
NaCl

What does the atomic number refer to? (Lit.)
The number of protons

How is the atomic mass determined? (Inf.)
Atomic mass is the total of the electrons, neutrons and protons of an atom.

Science (2)

If a good source of nectar has been discovered by a worker bee, how does the news get around the other foragers in his society? Bee language has been the subject of extensive study. A bee indicates the direction of food by dancing on the honeycomb. If the bee dances around and around, there is nectar somewhere within 75 yards. If the nectar is beyond this distance, the bee specifies the exact direction by repeatedly running in a straight line wagging its abdomen from side to side. More recent investigations report that not only does a bee tell other bees where to find food by dancing, but they pass on information by means of secretions from their bodies. Thus, worker bees lick a substance from the body of the queen and convey it from bee to bee in the food they share. By this means, thousands of bees in a colony are apprised of the presence of their queen although only ten or twenty of them are in direct contact with her at any one time. Any deficiency in the amount of the queen's substance is quickly detected by the colony as a whole and causes it to take steps to rear another queen. If a queen is removed from her colony, the workers become aware of her absence within an hour or two and excitedly search for her.

How soon do worker bees become aware of the queens absence? (Lit.)
Within two hours

How do bees share known sources of nectar with each other? (Inf.)
Repeated dances that tell the direction and distance to find nectar

What information besides source may be passed on from bee to bee? (Inf.)
Type of nectar or food available

How do bees know when to raise a new queen bee? (Lit.)
When the queen secretion is reduced to a certain point

How do beekeepers quiet hive work to move a hive? (BK)
They use smoke to force the guard bees inside

Science (3)
How reliable are measurements? Any measurement has some uncertainty associated with it, and the magnitude of that uncertainty is important. In football, the outcome of the game may depend on the crucial measurement of ten yards. The players rely on the precision with which the officials set the chain on the sidelines and the accuracy of the length of the chain. The terms precision and accuracy may mean the same thing to you but their meanings are very different in the sciences. Precision means the degree of reproducibility of a measurement. Accuracy means the degree to which a measurement represents the true value of what is measured. The precision of a measurement is determined by repeated individual measurements. For example, let's say you take your temperature three times in succession, and the readings are 98.6, 98.5, and 98.7 degrees Fahrenheit. Your temperature is best represented by the average of these readings, 98.6 degrees Fahrenheit. The temperature readings differ from the average value by about one tenth of a degree, which is the precision of the measurement.

What does the term precision refer to in science? (Lit.)
The degree of reproducibility of the measurement

How does accuracy differ from precision in measuring instruments? (Inf.)
Accuracy refers to how well it measures the "true" value, in addition to how reproducible it is

How are some scientific words misinterpreted in common language? (Inf.)
The meanings of some words sound similar to other words.

How does calibration assist in measuring accuracy? (BK)
Instruments are checked against a primary standard and adjusted with calibration

How would you determine the precision of a thermometer? (Lit.)
Take several readings and use the average of the readings

Science (4)

The Moon, Earth's only natural satellite, is large as far as moons go. It is fifth in diameter among planetary satellites, more than two thirds as large as Mercury, and more than three times the diameter of the largest asteroid. It is, in fact, over one fourth the size of the Earth with a diameter of 2,160 miles. Since the moon is a relatively near neighbor, we can measure its distance easily by geometrical methods. The moon travels around Earth in an elliptical orbit, with the Earth at one of the two focuses (foci) of the ellipse. As a result the distance to the moon varies by quite large amounts. Its close (or perigee) distance is 384,400 x (1- 0.055) = 363,258 km. Its far (or apogee) distance is 384,400 x (1+0.055) = 405,542 km. But these values turn out to be the mean or average values. In terms of miles the average is 238,857. Next to the Sun, the full moon is the brightest object in the heavens; however, its surface is rough and brownish and reflects light very poorly. In fact, the moon is about the poorest reflector in the solar system. The amount of light reflected by a celestial object is called the albedo. The moon reflects only seven percent of the sunlight that falls upon it so the albedo is 0.07.

In comparison to other planet's satellites, how does the moon compare in size? (Lit.)
It is larger than most

How is the albedo of a satellite determined? (Inf.)
This is the measure of the surface reflectivity of the sun's light

Why does the moon's distance from the earth vary? (Inf.)
It travels in an elliptical orbit around the earth

What is the name of a planet a little larger than the Moon? (Lit.)
Mercury

How long does it take the moon to orbit the earth? (BK)
About 28 days

Science (5)

Several generations of college students have grown up in a nuclear age. The atomic bombs that were used to devastate two Japanese cities and end World War II were small compared to the hydrogen bombs of today. This aspect of nuclear technology continues, as the major political powers try to limit and control their respective arms and prevent the spread of nuclear devices to smaller countries. Use of the nuclear energy of radio isotopes for the generating of power is a continuing and controversial story. The world is limited in its supply of fossil fuels and nuclear energy is currently being used as an alternative for the generation of electricity; however, there is much concern among Americans about the radioactivity hazards and environmental side-effects of nuclear power plants. The nuclear power industry in France is well supported as a matter of public policy. In contrast to the United States, France has continued to construct plants safely and with none of the delays experienced in this country. However critics argue that the French claim that nuclear power costs are "the lowest in the world" can't be substantiated because nobody knows the cost of the entire domestic nuclear program. For decades, the civilian program has profited from direct and indirect subsidies, in particular through cross-financing with the nuclear weapons program. Current estimates don't appropriately take into account eventual decommissioning and waste-management costs, which remain a concern and quite uncertain. (In addition to post-fission waste, 46 years of uranium mining has left 50 million tons of waste for eventual cleanup and remediation, the cost of which is unknown.) Official final disposal cost estimates for long-lived high- and intermediate level fission wastes vary between $21 billion and $90 billion.

How long have we been in the so called "nuclear age"? (Inf.)
Several decades (early 1940s)

What shortage of energy supply has prompted the interest in nuclear power? (Lit.)
Fossil fuels

What concerns do people in the United States have regarding nuclear power? (Inf.)
Damage to the environment and people's health

How do the French feel toward nuclear power? (Lit.)
Favorable for the most part

How is waste from nuclear power plants treated? (BK)
Buried in isolated regions, or kept in holding tanks

Science (6)

Igneous rocks are records of the thermal history of Earth. Their origin is closely associated with the movement of tectonic plates and they play an important role in the spreading of seafloor, the origin of mountains, and the evolution of continents. The best known examples of igneous activity are volcanic eruptions in which liquid rock material works its way to the surface and erupts from volcanic fissures and vents. Less obvious, although just as important are the enormous volumes of liquid rock that never reach the surface but remain trapped in the crust where they cool and solidify. Granite is the most common variety of igneous rock and is typically exposed in the eroded mountain belts and in the roots of ancient mountain systems now preserved in the shields. In January 1983, Kilauea volcano began an eruption that is still on-going. This volcano on the south shore of Hawaii is Earth's most active volcano. Volcanoes like Kilauea are dramatic proof that Earth's interior is still warm and active.

What is the most common variety of igneous rock? (Lit.)
Granite

What does the term "lithosphere" refer to? (BK)
The outer solid part of the earth, including the crust and mantle

What type of rock flows during volcanic eruptions? (Inf.)
Hot molten rock

How can scientists examine Granite? (Inf.)
In the field and in the lab

Where is the Kilauea volcano located? (Lit.)
Hawaii

History (1)

Jack Kennedy was the youngest American ever to be elected president of the United States. He was also the first Roman Catholic to hold that office. When he was assassinated in Dallas, Texas on November 22, 1963 halfway through his first term of office the entire nation mourned. Kennedy was the second son of Joseph Kennedy, a multimillionaire of Irish decent. After academic training, he served in World War II, and as a torpedo boat commander he was decorated for bravery. After the war ended Kennedy went into politics as a Democrat, and by 1960 had built up enough support to stand for the presidency. He won with a small majority, but what is interesting is that the largest share of his vote came from new voters. He was seen to represent the hopes and dreams of young men and women, especially colored people and they looked to him to build a new world. During his term he introduced some civil rights laws and he planned others. He handled the foreign policy of the United States with skill and courage, particularly in 1962, when he prevailed upon the Russians to withdraw missiles from Cuba by making it clear that he would not hesitate to use nuclear weapons if a war followed. Kennedy was murdered in 1963 by a single rifle bullet, but the identity of the assassin has never been absolutely established.

What religion was Jack Kennedy? (Lit.)
Roman Catholic

Who was the main group voting for Kennedy? (Inf.)
Younger voters

What year was Jack Kennedy assassinated? (Lit.)
1965

For what branch of service did Jack Kennedy serve? (Inf.)
The Navy

How many sons of Joseph and Rose Kennedy went into politics? (BK)
Three

History (2)

Lewis and Clark led an expedition sponsored by the United States government through the vast wilderness that now makes up the western United States. They started their journey in May, 1804 near St. Louis, Missouri and returned in September, 1806. Lewis and Clark made maps along the way of their route and the surrounding regions. They also included specimens and descriptions of plants, animals and mineral resources as well as information about the natives who lived in the west. This information made it possible for the United States to claim the Oregon region. The idea for the expedition came from President Thomas Jefferson after the United States purchased the Louisiana Territory from France. He believed the journey would help the United States to claim the Oregon region. Meriwether Lewis was chosen to lead the expedition; he was a US Army captain and Jefferson's private secretary. William Clark was chosen by Lewis to join him. Clark had excellent map making skills and Lewis had training in the study of animals and plants. The journey started off with about fifty men. Many of the men were French boatmen hired to move the heavy keelboat against the Missouri River's swift current. During the winter of 1804, the explorers camped near Fort Mandan. At this time they met Toussaint Charbonneau, a French-Canadian trader, and his wife, a Shoshone Indian. The two joined the expedition and Sacagawea was a great help.

On what river did Lewis and Clark begin their journey? (Inf.)
Missouri River

How many men started the expedition? (Lit.)
Fifty men started the journey

What was the name of the first tribe of Native Americans Lewis and Clark encountered? (BK)
The Yankton Sioux

Who sent Lewis and Clark on their exploration west? (Lit.)
President Thomas Jefferson

How did Sacagawea help Lewis and Clark? (Inf.)
She was a guide and interpreter

History (3)

The Hindenburg disaster was one of the world's most tragic occurrences ever. The Hindenburg was a rigid air ship built by a firm in Germany. In 1936, the Hindenburg was completed and tested. It was the world's first trans-Atlantic airliner with a length of 804 feet and an outmost diameter of 105 feet. The Hindenburg was kept overhead by over 200 cubic meters of Hydrogen in 16 cells. Hydrogen, the first element in the periodic table, is an invisible gas that contains a single proton and one outer electron. The ship had four 100 horsepower Diamler-Benz diesel engines that allowed the airship to travel at top speeds of 82 miles per hour. In May of 1936, it underwent the first scheduled air service across the Atlantic from Frankfurt, Germany and Lakehurst, New Jersey. More than 70 passengers were onboard the Hindenburg. This great airship had a library, a dining room, and an exquisite lounge. However, the Hindenburg was struck with disaster on May 6, 1937. The hydrogen of the airship was ignited while maneuvering to land at Lakehurst, New Jersey. The Hindenburg was destroyed by the fire caused by the invisible gas, and 35 passengers and crew died.

How fast could the Hindenburg travel? (Lit.)
82 miles per hour

What type of fuel is Hydrogen? (Inf.)
Highly flammable and invisible

Where did the Hindenburg travel? (Lit.)
Across the Atlantic from Germany to New Jersey

What types of passengers were generally on the Hindenburg? (Inf.)
Affluent mostly rich businessmen

How was the Hindenburg used by the Nazis? (BK)
As a propaganda device for supporting Hitler's re-election

History (4)

From a political point of view the Civil War did not end in 1865, nor did it end in 1877 when the North gave up trying to control the South by force. Indeed, the effects of the war are with us today more than a century after the confederacy collapsed. In the 1850s, the controversy of slavery in the territories led most white southerners to become Democrats. When the war ended most stayed Democrats. After southern whites regained control of their local governments in the 1870s, they voted Democrat in national elections almost to the man. With southern blacks not permitted to vote, the Republican Party had no chance at all in any southern states. People spoke of the "solid south." Every state that succeeded from the union cast its electoral votes for the Democratic candidate in every presidential election from 1890 until 1928. The Republican Party had become the leading party in the north and west by 1860. It remained so throughout the decades after the Civil War. Memories of the war stirred up strong emotions and had a great influence on how people voted. Tens of thousands saw the Democrats as the disloyal dividers of the United States. The view stayed long after slavery had been done away with, and the idea of succession abandoned by even its most extreme southern supporters.

What perspective is this passage written from? (Inf.)
From a bi-partisan political perspective

What political party was the leading party in the North and the West in 1865? (Lit.)
The Republican Party

How did northerners view the Republican Party? (Lit.)
As the saviors of the union

What was the position of the Southern Democrats regarding slavery in the 1800's? (Inf.)
They were pro-slavery

What was a vigilante organization formed by the Southern Democrats? (BK)
Ku Klux Klan and the White League

History (5)

By 1968, things had gone from bad to worse for the Johnson administration regarding the Vietnam War. In late January, the Democratic Republic of Vietnam (DRV) and National Liberation Front (NLF) launched coordinated attacks against major southern cities. These attacks, known in the west as the Tet Offensive, were designed to break the aggressive will of the Johnson administration and force Washington to the bargaining table. The Communist Party believed the American people were growing war weary and that Hanoi could humiliate Johnson and force a peace upon him. Most of Hanoi's predictions about the Tet Offensive proved elusive. Communist forces suffered tremendous casualties in the south and the massacre of thousands of non-communists in Hue, during the Tet Offensive created ill will among many Hanoi supporters. Furthermore, several leading southern generals thought the plans for the Tet Offensive too risky and this created a strain in the relations between northern and southern communists. In any event, in late March, 1968 a disgraced Lyndon Johnson announced he would not seek the Democratic Party's re-nomination for president and hinted that he would go to the bargaining table to end the war. Johnson engaged the Vietnamese in secret negotiations in the spring of 1968 in Paris, and it soon was made public that the Americans and Vietnamese were meeting to discuss a means to an end to the long and costly war. Despite the progress in Paris, the Democrats could not rescue the presidency from Republican challenger Richard M. Nixon who claimed he had a secret plan to end the war.

Who organized the Tet Offensive? (Lit.)
The North Vietnamese military

What was Lyndon Johnson position on the Vietnam War? (Inf.)
He was in support of the Viet Nam War

Who was the president when the Vietnam War began? (BK)
John F. Kennedy

Where did Johnson meet secretly with the Vietnamese? (Lit.)
Paris

Why did the Republican Party win in 1968? (Inf.)
Americans were angry about the war

History (6)

In the 1930s, the entire world economy was in a slump. It was one of the worst depressions the world had known. America, a world leader, was suffering along with everyone else. The depression had devastating effects on the country, the stock market was in shambles, banks everywhere went under, business could not continue to operate, and farmers fell into bankruptcy. A quarter of the working force, or 13 million people were unemployed in 1932 and this was only the beginning. During the previous decade, the nation enjoyed a seemingly endless period of prosperity. The Great War ended in 1918, and then following a post war depression in 1920 and 1921 the economy took off. Under the leadership of President Harding, taxes were cut and so was spending. The president then retreated to his closet with his mistress leaving the economy to its own devices. Unfortunately for Harding, the strain of keeping his wife, "the duchess" away from his mistress became too much for him. In 1923, vice president, Calvin "Silent Cal" Coolidge took Harding's place as president. Coolidge was a tight lipped but popular president, and it was a good thing too. Unfortunately for Coolidge there was a lot more going on behind the scenes, besides Harding's shenanigans. The scandals of the Harding administration, including the "Tea Pot Dome," became public. The job of restoring faith in the government rested on his shoulders.

How many people were unemployed in 1932? (Lit.)
13 million, or one quarter of the working force

What problems did President Harding have during his administration? (Inf.)
Scandals about a mistress and bribes

What happened to farmers during the great depression? (Lit.)
Many went bankrupt

Why did Calvin Coolidge take over as president? (Inf.)
Harding died in office

What was the scandal involving the Tea Pot Dome? (BK)
Harding's secretary of the interior got a kickback for oil leases

REFERENCES

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. <u>Journal of Experimental Psychology: General, 131</u>(4), 567-589.

Afflerbach, P. (1986). The influence of prior knowledge on expert readers' important assignment process. In J. A. Niles & R. V. Lalik (Eds.). <u>National Reading Conference Yearbook. 35. Solving problems in literacy: Learners, teachers and researchers</u> (pp. 30-40). Rochester, NY: National Reading Conference.

Afflerbach, P. (2008). "National Reading Conference Policy Brief: High Stakes Testing and Reading Assessment". <u>Journal of Literacy Research, 37</u>(2), 151-162.

Anderson, R. C. (1978). How to construct achievement tests to assess comprehension. <u>Review of Educational Research, 42,</u> 145-170.

Beck, I. L., McKeown, M. G., & Kucan, L. (2005). Choosing words to teach. In E. H. Hiebert & M. L. Kamil (Eds.).<u>Teaching and learning vocabulary: Bringing research to practice.</u> Mahwah, NJ: Lawrence Erlbaum Associates.

Bell, P., & Linn, M. C. (2002). Beliefs about science: How does science contribute? In B. K. Hofer & P. R. Pintrich (Eds.) <u>Personal epistemology: The psychology of beliefs about knowledge and knowing</u> (pp. 321-346) Mahwah, NJ: Erlbaum.

Bennet, R. E. (1998). Reinventing assessment: Speculations on the future of large-scale educational testing. A policy information perspective. <u>Policy Information Center.</u> Princeton, NJ: Educational Testing Services. Available: http://www.ets.org.

Brown, J. I., Fishco, V. V., & Hanna, G., (1993) <u>Nelson-Denny Reading Test.</u> Itasca, IL: Riverside Publishing Company.

Bowyer-Crane, C., & Snowling, M. J. (2005). Assessing children's inference generation: What do tests of reading comprehension measure? <u>British Journal of Educational Psychology, 75,</u> 189-201.

Burton, C., & Daneman, M. (2007). Compensating for a limited working memory capacity during reading: Evidence from eye movements. <u>Reading Psychology, 98,</u> 163-187

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skill. Journal of Educational Psychology, 96, 31-42

Campbell, J. R. (1999). Cognitive processes elicited by multiple-choice and constructed response questions on an assessment of reading comprehension. Unpublished doctoral dissertation. Temple University.

Chi, M. T. H., Feltovitch, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

Cohen, P. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. Journal of Verbal Learning & Verbal Behavior, 19(4), 450-466.

Daneman, M., & Hannon, B. (2001). Using working memory theory to investigate the construct validity of multiple-choice comprehension on the SAT. Journal of Experimental Psychology: General, 130, 208-223.

Daneman, M., & Merikle, P. M. (1996). Working memory, and language comprehension: A meta-analysis. Psychometric Bulletin & Review, 3(4), 422-433.

Davis, F. B. (1942). Two measures of reading ability. The Journal of Educational Psychology, 33(5), 365-372.

Deane, P., Sheehan, K. M., Sabatini, J., Futagi, Y., & Kostin, I. (2006). Differences in text structure and its implications for assessment of struggling readers. Scientific Studies of Reading, 10(3), 257-275.

DeBacker, T. K., Beesley, A. D., Thoma, S. J., & Hestevold, N. L. (2008). The challenge of measuring epistemic beliefs: An analysis of three self-report instruments, The Journal of Experimental Education, 76(3), 281-312.

Duran, N. D., McCarthy, P. M., Graesser, A. C., & McNamara, D. S. (2007). Using temporal cohesion to predict temporal *coherence* in narrative and expository texts. Behavior Research Methods, 39(2), 212-223.

Embretson, K. A., & Wetzel, C. D., (1987). Component latent trait models for paragraph comprehension. Applied Psychological Measurement 11, 175-193.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. Journal of Experimental Psychology, 128(3), 309-331.

Everson, H. T., Osterlind, S. J., Dogan, E., & Tirre, W. (2007). The performance effects of word locator cues on the NAEP reading assessment. Practical Assessment, Research & Evaluation, 12(13). Retrieved from http//pareonline.net /getvn.asp?v=12&n=13

Garnham, A. (1981). Mental models as representations in text. Memory & Cognition, 9, 560-565.

Gorin, J. S. (2005). Manipulating processing difficulty of reading comprehension questions: The feasibility of verbal item generation. Journal of Educational Mesurement, 42(4), 351-373

Gorin, J. S., Embretson, K. A., & Sheehan, K. (2002). Cognitive and psychometric modeling of text-based reading comprehension GRE-V items, Paper presented at an annual meeting of the National Council on Measurement in Education, New Orleans, LA.

Hannon, B., & Daneman, M. (2006). What do tests of reading comprehension ability such as the VSAT really measure? A componential analysis. In: A. V. Mittel (Ed.). Focus on Educational Psychology. (pp. 105-146). Hauppauge, NY: Nova Science Publishers.

Johnston, P., & Afflerbach, P. (1982). The process of constructing main ideas from text. Cognition and Instruction, 2, 207-232.

Kardash, C. M., & Scholes, R. J. (1996). Effects of preexisting beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. Journal of Educational Psychology, 88(2), 260-271.

Katz, S., Lautenschalger, G. J., Blackburn, A. B., & Harris, F. H. (1990). Answering reading comprehension questions without passages on the SAT. Educational and Psychological Science 1, 122-127.

King, P. M., & Kitchener, K. S. (2002). The Reflective Judgment model: Twenty years of research on epistemic cognition. In B. K. Hofer & P. R. Pintrich (Eds.) Personal epistemology: The psychology of beliefs about knowledge and knowing (pp. 37-62), Mahwah: NJ: Erlbaum.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction integration model. Psychological Review, 95(2), 163-182.

Kintsch, W. (1998). Camprehension: A paradigm for cognition. New York: Cambridge University Press.

Kyllonen, P. C., & Cristal, R. E. (1990). Reasoning ability is (little more than) working memory capacity. Intelligence, 14(4), 389-433.

Lundeberg, M. (1987). Metacognitive aspects of reading comprehension: Studying understanding in legal case analysis. Reading Research Quarterly, 22, 407-432.

MacGinite, W.H., MacGinite, R. K., Maria, K., & Dreyer, L. G., (2002). Gates-MacGinite Reading Tests, 4th Edition, Forms S and T. Itasca, IL: Riverside Publishing Company.

Magliano, J. P., & Millis, K. K. (2003). Assessing reading skills with a think-aloud procedure and latent sematic analysis. Cognition and Instruction, 21, 251-284.

Magliano, J. P., & Radvansky, G. A. (2001). Goal coordination in narrative comprehension. Psychonomic Bulletin & Review, 8(2), 372-376.

Manhart, J. L. (1996). Factor analytic methods for determining whether multiple-choice and constructed-response tests measure the same construct. Paper presented at the Annual Meeting of the National Council on Measurement (NY, April 9-11)

Mason, L., Gava, M., & Boldrin, A. (2008). On warm conceptual change: The interplay of text, epistemological beliefs, and topic interest. Journal of Educational Psychology, 100(1), 297-309.

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

McNamara, D., Kintsch, W., Songer, N. B., & Kintsch, E. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and level of understanding in learning from text. Cognition and Instruction, 82, 1-43.

McNamara, D., & McDaniel, M. A. (2004). Suppressing irrelevant information: Knowledge activation or inhibition? Journal of Experimental Psychology: Learning, Memory & Cognition, 30, 465-482.

Millis, K., Magliano, J., & Todaro, S. (2006). Measuring discourse-level processes with verbal protocols and latent semantic analysis. Scientific Studies of Reading, 10(3), 225-240.

Moss, B. (2008). The informational text gap: The mismatch between non-narrative text types in basal readers and 2009 NAEP recommended guidelines. Journal of Literature Research, 40(2), 201-219.

Mulligan, E. J. (2006). Assessment on individual differences in the components of comprehension skill. Unpublished doctoral dissertation. University of Colorado.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse-representation during reading. Discourse Processes, 26, 131-157.

National Institute of Child, Health and Human Development (NICHD), (2002). Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. (Reports of the Subgroups). Washington D.C.: US Government Printing Office.

O'Brien, E. J., & Myers, J. L. (1999). Text comprehension: A view from the bottom up. In S. R. Goldman, A. C. Greasser & P. van den Broek (Eds.). Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso (pp. 35-53). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Ozuru, Y., Best, R., Bell, C., Witherspoon, A., & McNamara, D. S. (2007). Influence of question format and text availability on assessment of expository text comprehension. Cognition and Instruction, 25(4), 399-438.

Pearson, P. D., & Hamm, D. N., (2005). The assessment of reading comprehension: A review of practices—past, present and future. In S. G. Paris & S. A. Stahl (Eds.) Children's reading comprehension and assessment (pp. 13-69). Mahwah, NJ: Erlbaum.

Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2007). Vocabulary assessment: What we know and what we need to learn. Reading Research Quarterly, 42(2), 282-296.

Perfetti, C. (1997). Sentences, individual differences, and multiple texts: Three issues in text comprehension. Discourse Processes, 23, 337-55.

Perfetti, C., & Lesgold A. M. (1977). Discourse comprehension and sources of individual differences, In M. A. Just & P. A. Carpenter (Eds.) Cognitive processes in comprehension (pp. 141-183). Hillsdale, NJ: Erlbaum.

Perry, W. G., Jr. (1970). Forms of intellectual and ethical development in the college years: A scheme. New York: Holt Rinehart and Winston.

Pintrich, P. (2002). Future challenges and directions for theory and research on personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.) Personal epistemology: The psychology of beliefs about knowledge and knowing. (pp. 389-414). Mahwah, NJ: Erlbaum.

Radvansky, G. A., & Copeland, D. E. (2004). Working memory span and situation model processing. The American Journal of Psychology, 117(2), 191-213.

Rukavina, I., & Daneman, M (1996). Integration and its effect on acquiring knowledge about competing scientific theories from text. Journal of Educational Psychology, 88(2), 272-297.

Ryan, M. P. (1984). Conceptions of prose coherence: Individual differences in epistemological standards. Journal of Educational Psychology, 76(6), 1226-1238.

Sarroub, L., & Pearson, P. D. (1998). Two steps forward and three steps back: The story history of reading comprehension. Clearing House, 72(2), 97-105. Journal code CIJAUG1999.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. Journal of Educational Psychology, 82, 498-504.

Schommer-Aikins, M. (2004). Explaining the epistemological belief system: Introducing the embedded systemic model and coordinated research approach. Educational Psychologist, 39(1), 19-29.

Schneider, W., Eschman, A., & Zuccolotto, A., (2002). E-Prime user's guide (Version 1.1). Pittsburg Psychology Software Tools Inc.

Schraw, G., Bendixen, L. D., & Dunkle, M. E. (2002). Development and validation of the Epistemic Beliefs Inventory (EBI). In B. K. Hofer & P. R. Pintrich (Eds.) Personal epistemology: The psychology of beliefs about knowledge and knowing. (pp. 261-275). Mahwah, NJ: Erlbaum.

Shapiro, A. M. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. American Educational Research Journal, 41, 159-189.

Sheehan, K. M., & Ginther, A. (2001). What do passage-based multiple choice verbal reasoning items really measure? An analysis of the cognitive skills underlying performance on the current TOEFL reading section. Paper presented at the 2000 Annual Meeting of the National Council of Measurement in Education, New Orleans, LA.

Singer, M. (1988). Inferences in reading comprehension. Reading research: Advances in theory and practice. New York: Academic Press.

Singer, M (1994). Discourse inference processing. Handbook of psycholinguistics. New York: Academic Press.

Singer, M., & O'Connell, G. (2003). Robust inference processes in expository text comprehension. European Journal of Cognitive Psychology, 15(4), 607-631.

Smith, M. (1991). Put to the test: The effects of external testing on teachers. Educational Researcher, 20, 8-11.

Tabachnick, B. G., & Fidell, L. S., (2007). Using multivariate statistics 5th Ed. Boston, MA: Allyn and Bacon

Valencia, S., & Pearson, P. D. (1987). Reading assessment: Time for a change. The Reading Teacher, 40, 726-732.

Van Dijk, T. A., & Kintsch, W. (1983). Strategies for discourse. New York: Academic Press.

Voss, J., & Silfies, L. (1996). Learning from history text: The interaction of knowledge and comprehension skill with text structure. Cognition and Instruction, 14, 45-68.

Utah State Office of Education (2002). Utah currently uses the Utah Basic Skills Competency Test. Retrieved 5/5/2008, www.schools.utah.gov/assessment/documents/UBSCT.

Was, C. A. (2005). Multiple automatic and controlled working memory processes in comprehension: A latent variable approach. Unpublished doctoral dissertation. University of Utah.

Was, C. A., & Woltz, D. (2007). Reexamining the relationship between working memory and comprehension: The role of available long-term memory. Journal of Memory and Language, 56, 86-102.

Whitney, P., Arnett, P. A., Driver, A., & Budd, D. (2001). Measuring central executive functioning: What's in a reading span. Brain and Cognition, 45, 1-14.

Wiley, J., & Myers, J. L. (2003). Availability and accessibility of information and causal inferences from scientific text. Discourse Processes, 36(2), 109-129.

Woltz, D. (1988). An investigation in the role of working memory in procedural skill acquisition. Journal of Experimental Psychology: General, 117, 319-331.

Wood, P., & Kardash, C., (2002). Critical elements in the design and analysis of students if epistemology. In B. K. Hofer & P. R. Pintrich (Eds.) Personal epistemology: The psychology and beliefs about knowledge and knowing. (pp. 231-260). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.