THE CATHOLIC UNIVERSITY OF AMERICA

An Investigation of the Effects of Reader Characteristics on Reading Comprehension Of a General Chemistry Text

A DISSERTATION

Submitted to the Faculty of the

Department of Education

School of Arts and Sciences

Of The Catholic University of America

In Partial Fulfillment of the Requirements

For the Degree

Doctor of Philosophy

By

Kelly Y. Neiles

Washington, D.C.

2012

An Investigation of the Effects of Reader Characteristics on Reading Comprehension
Of a General Chemistry Text

Kelly Y. Neiles, PhD.

Director: Dr. Diane M. Bunce, PhD.

There is great concern in the scientific community that students in the United States, when compared with other countries, are falling behind in their scientific achievement. Increasing students' reading comprehension of scientific text may be one of the components involved in students' science achievement. To investigate students' reading comprehension this quantitative study examined the effects of different reader characteristics, namely, students' logical reasoning ability, factual chemistry knowledge, working memory capacity, and schema of the chemistry concepts, on reading comprehension of a chemistry text. Students' reading comprehension was measured through their ability to encode the text, access the meanings of words (lexical access), make bridging and elaborative inferences, and integrate the text with their existing schemas to make a lasting mental representation of the text (situational model). Students completed a series of tasks that measured the reader characteristic and reading comprehension variables. Some of the variables were measured using new technologies and software to investigate different cognitive processes. These technologies and software included eye tracking to investigate students' lexical accessing and a Pathfinder program to investigate students' schema of the chemistry concepts.

The results from this study were analyzed using canonical correlation and regression analysis. The canonical correlation analysis allows for the ten variables described previously to be included in one multivariate analysis. Results indicate that the relationship between the

reader characteristic variables and the reading comprehension variables is significant. The resulting canonical function accounts for a greater amount of variance in students' responses then any individual variable. Regression analysis was used to further investigate which reader characteristic variables accounted for the differences in students' responses for each reading comprehension variable. The results from this regression analysis indicated that the two schema measures (measured by the Pathfinder program) accounted for the greatest amount of variance in four of the reading comprehension variables (encoding the text, bridging and elaborative inferences, and delayed recall of a general summary). This research suggest that providing students with background information on chemistry concepts prior to having them read the text may result in better understanding and more effective incorporation of the chemistry concepts into their schema.

This dissertation by Kelly Y. Neiles fulfills the dissertation requirement for the doctoral degree in educational psychology approved by Dr. Diane M. Bunce, PhD., as Director, and by Dr. Kathleen C. Perencevich, PhD., and Dr. John J. Convey, PhD. as Readers.

Dr. Diane M. Bunce, PhD., Director

Dr. Kathleen C. Perencevich, PhD., Reader

Dr. John J. Convey, PhD., Reader

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

.

Dedication

I dedicate this Doctoral dissertation to my
Mom, Dad, Brady, and Aunt Peg
Whose continued support and words of
Encouragement allowed me to complete this process.

Acknowledgements

Dr. Diane M. Bunce—The Catholic University of America
Dr. Kathleen C. Perencevich—The Catholic University of America
Dr. John J. Convey—The Catholic University of America
Elizabeth A. Flens—The Catholic University of America
Dr. Frank R. Yekovich—The Catholic University of America

CHAPTER ONE

Introduction to the Problem

There is great concern in the scientific community that students in the United States, when compared to other countries, are falling behind in their performance (National Research Council, 2007). Reading comprehension of scientific text, scientific literacy, may be one of the critical components for increased science achievement (Otero, Leon, & Graesser, 2002; Ozuru, Dempsey, & McNamara, 2009). Currently, research in scientific literacy has not been extensive enough to provide a full understanding of the processes involved in reading comprehension when the text is scientific in nature. A small number of studies have shown large correlations between reading comprehension and science proficiency (Cromley, 2009; O'Reilly & McNamara, 2007). While these studies are promising, they are only the start of the research that is needed. A better understanding of the comprehension of science text by students may improve their science achievement in the future (Cromley & Snyder-Hogan, 2010).

Few studies have been conducted to investigate scientific literacy when the content in the text is chemistry. A search of research literature results in a handful of studies that have investigated the cognitive processes involved in the comprehension of chemistry text (Bond, 1941; Pellettieri, 1955; Schmuckler, 1983; Wilson & Neubauer, 1988). These studies are focused on the type of text that is found in chemistry text books and the presentation of the information, not the students' interaction with the information. This focus on the type of text is troubling since chemistry text is often characterized by explanatory text structure and heavy vocabulary demands (Otero, Leon, & Graesser, 2002). These components increase the complexity of the text and make the processes involved in comprehension more extensive. More studies investigating the cognitive processes involved in the comprehension of chemistry text are needed so that we may better understand these processes. Better understanding of the processes may lead to better instruction resulting in an increase in students' achievement in chemistry.

While an increase in the students' achievement is the ultimate goal, we must first understand the processes involved in reading comprehension. This study focused on the effects of reader characteristics including student schema, logical reasoning, working memory, and factual knowledge on reading comprehension when students were presented with a passage from a general chemistry text. The measures of reading comprehension included the students' ability to encode the textual information, access relevant lexical information, make necessary inferences, and create long term situational models using the textual information given. The study relied on various measures to investigate each of these processes including recall of text, eye tracking protocol while reading the text, multiple choice inference questions answered after the text has been read, and delayed recall of the general information in the text within 24 hours after the initial reading. This chapter provides an overview of the background information, significance of the study, and methodology used. The chapter will conclude by identifying the research questions investigated along with the predicted hypotheses.

Scientific Literacy and Scientific Text

The role of literacy in the teaching and learning of science has gone through many changes in the past few decades (Norris & Phillips, 2003). The traditional view of reading in science is one of a passive process in which the students are using the text only to get to the scientific information it holds (Wellington & Osborne, 2001). This view of reading has created an environment where science teachers have little concern for the text and view reading as an unimportant part of science education. A perception of scientific text as only a tool to get at the scientific information can be partly attributed to the fact that many science teachers see the process of reading a scientific text as a simple process (Norris & Phillips, 2003). Teachers are well versed in the scientific terms used in the text as well as the layout of the text so they can easily understand most scientific text presented. This is not necessarily true of an inexperienced student. Students known as 'good' readers are usually those students who know the words, read flexibly, identify and locate information, and recall content well (Haas & Flower, 1988). When these students are presented with the task of analyzing, criticizing, or interpreting the text, they cannot do so. A traditional view of good reading is not adequate to explain what is expected of students when they read scientific text. The act of reading a scientific text involves an in depth interaction with the text (Norris & Phillips, 2003). This disconnect between traditionally good readers and good scientific readers necessitates a two pronged theory of scientific literacy. Norris and Phillips (2003) describe a two pronged theory that includes the ability to read and write when the content is science (*fundamental* sense), and being knowledgeable, learned, and educated in science (*derived* sense).

Fundamental and Derived Senses of Scientific Literacy

For a student to be scientifically literate he/she must be successful in both the *fundamental* sense of literacy and the *derived* sense of literacy (Norris & Phillips, 2003).

Many science classes only emphasize the second of these, the *derived sense* (Pressley & Wharton-McDonald, 1997). The *derived* sense of literacy involves learning the facts and concepts involved in a scientific field (Norris & Phillips, 2003). This sense of literacy can be thought of as the ability to understand the concepts presented in the text. For example, a chemistry student who is scientifically literate in the *derived* sense will have an understanding of the various chemical processes involved in the passage he/she is reading (chemical equilibria, reactions, etc.). Often a student's understanding of these concepts is taken as being scientifically literate (Norris & Phillips, 2003). This may not be the case since the student can understand the underlying concepts and effectively read the text but may not be able to critically evaluate or interpret the text containing these concepts. This ability to critically evaluate or interpret is called the *fundamental* sense of scientific literacy (Norris & Phillips, 2003). In this sense the student is able to determine the credibility and/or usefulness of the text. This sense of literacy involves inferring meaning from the text and integrating it with what the student already knows. The result is something greater than the text itself or the student's knowledge; it is the student's interpretation of the text (Phillips, 2002). This interpretation goes beyond what is in the text and will not necessarily occur just because the student is traditionally a 'good' reader.

The necessity of scientific literacy in both the *fundamental* and *derived* sense is important to the teaching of science (Norris & Phillips, 2003). Scientific literacy can be different from other literacies due to the difficult content involved (Otero, Leon, & Graesser, 2002). The scientific community has seen that focusing on the content specific to each science field has resulted in an almost exclusive focus on the substantive content in scientific literacy (Norris & Phillips, 2003). Focusing on content neglects the interpretative capacities required to successfully read the text. It has also resulted in many students having isolated understandings of the laws, facts, and theories in only one field of science (chemistry, biology, physics, etc.) without an understanding of the interconnections present between the sciences (Norris, 1992; Norris & Phillips, 2003). The comprehension, interpretive, analytical, and critical capacities needed for reading comprehension are largely the same for most substantive texts (Norris & Phillips, 2003). This implies that the educational goal of scientific literacy is common with the literacy goals in many other substantive content areas. By fostering literacy in both its *derived* and *fundamental* sense, we may be able to improve a students' literacy in many different areas of their science learning.

For the *fundamental* sense of scientific literacy students are not only learning the concepts presented in the text through decoding the words and locating information in them (as described in a traditional proficiency in reading), but must also develop the ability to read text in an interactive way that requires engagement with the text on the part of the student (Haas & Flower, 1988). To be successful in science, the student needs to be instructed not only in the substantive knowledge involved in the concept, but also in the generalizable concepts, skills, and understandings necessary for reading (Norris, 1989; Norris, 1992). Learning these skills can help the student in the field in which these skills were taught, as well as in other fields where the reading of substantive text is required.

To understand how to teach scientific literacy in the *fundamental* sense, first we must understand the different cognitive processes that enable reading comprehension. This includes the cognitive architecture of the mind and how the student integrates the incoming information with his/her prior knowledge.

Cognitive Architecture

The theory of cognitive architecture that will be used to describe the processes necessary for reading comprehension in this study is Anderson's (2007) Adaptive Control of Thought-Rational (ACT-R). This theory describes how the biological components of the human brain can achieve the various cognitive functions of the mind. It includes a description of the mind as modularly organized with each module having a specific function in cognitive processes. The theory also explains how these processes can work in parallel and/or together to achieve the overall functioning of the mind (Anderson, 2005). By understanding these underlying components in the functioning of the mind we can better understand some of the more complex cognitive processes that integrate these components (Anderson, 2007). One of these complex processes is the storage and retrieval of information in the mind described through the theory of schema (Mayer, 1983). The theory of schema takes the various components described in ACT-R and applies it to the complex organization of one's prior knowledge and the integration of novel information.

Schema

The theory of schema describes a person's organization of concepts as a general framework for the understanding of incoming information (Mayer, 1983). It is a general theory that describes a person's ability to store and retrieve knowledge. Because the theory of schema is general, it can be applied directly to many processes including reading and reading comprehension of text. Chomsky (2006) describes the theory of schema in reading through three basic ideas. The first is a distinction between the way a sentence is written (surface structure) and the way it is represented in memory (deep structure). The second is a set of rules for conversion between the surface structure of the text and the deep structure stored in memory. The final idea is that there are general characteristics that are shared by all language users in the ways that they acquire language. A reader will use his/her schema when learning information from reading a text (Bower, Black, & Turner, 1979; Mayer, 1983). This process can involve making changes to his/her existing schema. The schema can then be used in future situations when the information stored is relevant to new incoming information. Through this process a reader's prior knowledge in the content relevant to the text will affect his/her ability to comprehend and interpret that text (Bransford & Johnson, 1972; Pichert & Anderson, 1977; Schallert, 1976). This means that the reader's schema may have a significant effect on the different processes involved in reading comprehension.

Of the many reader characteristics that affect reading comprehension, a reader's schema may best explain the differences in the processes of reading comprehension (Mayer, 1983). The theory of schema explains how people store and use information as general ideas rather than as individual pieces of knowledge. This is evident in research that shows people tend to remember the gist of a passage rather than the verbatim content; important information better than unimportant information; and information that is consistent with their personal perspective better than information that is not (Mayer, 1983). The reader's schema affects many components of reading comprehension including the way he/she encodes text, accesses lexical information, makes inferences, and remembers the general ideas of a text that are integrated into the reader's existing schema.

Reading Comprehension

Reading comprehension involves two initial stages where the reader is internalizing the incoming textual information and creating a meaningful representation for himself/herself (Just & Carpenter, 1987). The first stage is the encoding of the text where the reader combines features of incoming stimuli to determine what words and sentences are present in the text (Gibson, 1969). The second stage is the process by which words are combined to form mental representations of the meaning of the text (Kintsch, 1974). This involves both the meanings of the words themselves, as well as the relationships found between the words. These relationships can be explicitly present in the text or derived from inferences made by the reader. Both of these processes, encoding and creation of a mental representation, are necessary for the reader to comprehend the text and must occur before a third process, utilization, can proceed (Just & Carpenter, 1987). The importance of the first two processes in reading comprehension and thus scientific literacy was the focus of this study.

Levels of Representation of Text

Reading comprehension and the processes involved can be investigated through the three levels of representation of text formed by the reader (Kintsch, 1998). These levels include the surface level, propositional level, and situational model. The surface level includes the encoding by the reader of the text presented. The surface level storage is of the exact verbiage used in the text itself (van Dijk & Kintsch, 1983). The propositional level is an abstraction of this verbiage in which the student represents the connections between the words but not the actual words themselves (Kintsch & van Dijk, 1978). The situational model is made up of the major ideas involved in the passage. It can be thought of as a summary of the text (Kintsch, 1998). The creation of these three levels of representation does not necessarily ensure that the representations are correct or without misconceptions (Just & Carpenter, 1987). Each level is a key part of the comprehension process for the student.

By having multiple levels of representation the reader is more likely to be able to utilize the information in working memory or, if necessary, commit the information to long term memory (Kintsch, 1998). If a reader successfully represents the text at all three levels, he/she will have a better comprehension of the material presented. All three levels will not necessarily be stored in the reader's memory indefinitely. In fact, only the situational model is likely to be stored in long term memory after the surface and propositional levels have been forgotten (Kintsch, 1998). By reducing the information stored in memory, the reader is maximizing the efficiency of this storage process.

The purpose of this study was to investigate how students' reader characteristics affect their ability to create the three levels of representation involved in reading comprehension. Five variables used to measure the three levels of representation included the following: encoding of text (surface level representation), lexical accessing of word meanings (propositional level representation), development of bridging and elaborative inferences (situational model), and recall of a general summary of the text (situational model). The investigation was conducted in the context of reading general chemistry passages. The reader characteristic variables consisted of two measures of students' schema in two general chemistry concepts as well as logical reasoning ability, factual chemistry knowledge, and working memory capacity. A list of the variables used in the canonical correlation analysis used in this study can be found in Table 1.

Table 1

Reader Characteristic Set and Reading Comprehension Set of Variables

Reader Characteristic Set	**Reading Comprehension Set**
Schema Measure One	Encoding
Schema Measure Two	Lexical Access
Logical Reasoning Ability	Bridging Inferences
Factual Chemistry Knowledge	Elaborative Inferences
Working Memory	Recall of General Summary of Text

The results from this canonical correlation will provide information about the cognitive processes involved in reading chemistry text. By identifying the reading comprehension processes most affected by a student's reader characteristics, we may be able to target instruction for these processes. This could result in increased reading comprehension for science-based texts. Because scientific literacy can be generalized across different science fields, this study may be indicative of the effects of reader characteristics on reading comprehension in scientific fields other than chemistry. An overview of the procedures used and variables chosen to investigate the levels of representation present in this study will be discussed in the following Methods section.

Methods

This research was a novel study in the integration of Just and Carpenter's (1987) theory of reading comprehension with Kintsch's (1998) levels of representation and other reader characteristics using a canonical correlation to determine relationships between the components. The reader characteristics evaluated in this study included the students' schema, logical reasoning, factual chemistry knowledge, and working memory (Bransford & Johnson, 1972; Cain & Oakhill, 2004; Chomsky, 2006; Ericsson & Kintsch, 1995; Palladino, Cornoldi, De Beni, & Pazzaglia, 2001). This new integrated theory was used to evaluate students' reading comprehension of a general chemistry text. Different quantitative measurements were used for each of the variables involved in the levels of representation in reading a scientific text as well as for the reader characteristic variables. All variables, reader characteristic and reading comprehension, were analyzed using a canonical correlation. The method for measurement of each variable was chosen based on the current research and practices utilized in the study of that variable. The operational definitions used in this study are given in Table 2.

Table 2

Operational Definitions

Term	**Operational Definition**
Schema	The general structure in which a person stores knowledge in long term memory.
Path Length Correlation	The similarity of the number and length of node to node paths within a student's schema of a subject when compared with the schema of an expert in that subject.
Neighborhood Similarity	The similarity of surrounding nodes around a specific node in a student's schema when compared with those of experts in that topic.
Working Memory	An active platform for processing incoming information with information from long term memory.
Factual Chemistry Knowledge	The main ideas of a chemistry topic.
Logical Reasoning Ability	The ability to determine whether something is true through inductive and deductive reasoning.
Encoding	Perceiving the printed symbols as a word.
Lexical Access	Accessing the word's meaning and relevant relationships in the context of the text.
Bridging Inference	The act of recognizing connections between what was most recently read to some portion of the earlier text.
Elaborative Inference	The act of creating connections between what was read in the text with the student's prior knowledge.
General Summary of Text	The main idea of a passage and other pieces of important information ordered hierarchically.
General chemistry text	The written words of a paragraph when the content is chemical in nature and written at an introductory level.

This study involved two sets of continuous variables. The first set is described as the reader characteristic set and include the following five variables: logical reasoning ability, factual chemistry knowledge in the relevant concepts, working memory, and two measures of schema (path length and neighborhood content which will be discussed in further detail in Chapter 2). The second set of variables is the reading comprehension set and includes five variables: encoding, lexical access, bridging inferences, elaborative inferences, and the ability to remember a general summary of the text. Because there are multiple variables in each variable set and all variables are continuous, a canonical correlation was used for the statistical analysis (Stevens, 2009). This analysis is the multivariate generalization of the Pearson Product Moment Correlation. It determines the weights or values assigned to each variable so that a correlation using the linear combination of each set of variables (reader characteristic set and reading comprehension set) is as highly correlated as possible. These weights can be used to evaluate the amount of variance each reader characteristic variable accounts for in the reading comprehension variables.

The results from this analysis will indicate how the variables in the reader characteristic set are related to those in the reading comprehension set. It will also determine which reader characteristic variables contribute significantly to the prediction of the reading comprehension variables. It will likewise determine which reader characteristic variables are better predictors for the reading comprehension set as a whole. The analysis will provide separate regressions for each of the five reading comprehension variables. Though these separate regressions are not as statistically strong as the canonical correlation, they will provide information as to how the reader characteristic variables affect each reading comprehension variable.

Measurements of Reader Characteristic Set

Logical Reasoning Ability, Factual Chemistry Knowledge, and Working Memory

Students' logical reasoning, factual chemistry knowledge, and working memory were measured in this study as three of the five variables in the reader characteristic set (along with two measures of schema). Logical reasoning has been shown to affect reading comprehension (Spearritt, 1972). In this study logical reasoning ability was measured using the Group Assessment of Logical Thinking (GALT) test constructed by Roadrangka, Yeany, and Padilla (1983). The GALT test has been used reliably in general chemistry classes as a measure of logical reasoning ability and has also been shown to be a good predictor of chemistry achievement (Bunce & Hutchinson, 1993). Each student received a score of 1 (low logical reasoning ability) to 12 (high logical reasoning ability) on this test. The GALT test used in this study is found in Appendix A.

Students' knowledge in a topic area can also affect their reading comprehension when the text content is based on that subject (Cain & Oakhill, 2004). In this study factual chemistry knowledge was considered different from schema as it involved isolated facts related to the topics. Unlike schema, factual knowledge does not involve the relationships between the different concepts within the topic (which was measured as schema in this study). Factual chemistry knowledge was measured using a multiple choice chemistry exam developed and validated by the researcher (Appendix B). This test covered the main ideas or facts involved in each of the chemistry topics used in the student portion of this study. The students received a score from 0 (low factual knowledge) to 20 (high factual knowledge) on this test.

Working memory capacity has also been shown to have an effect on reading comprehension (Cain & Oakhill, 2004; Ericsson & Kintsch, 1995). Reading comprehension involves many components that need to be stored in working memory including sensory features, linguistic expressions, propositional structures, and situational models (van Dijk & Kintsch, 1983). The necessity for various components in the working memory leads to the reader needing sufficient storage capacity in his/her working memory so that a reader can accommodate all necessary components. Working memory in this study was investigated using an electronic digit span test. The students received a score that corresponds with the number of items they were able to successfully hold in working memory from 0 (no items held in working memory) to the maximum number of items they were able to hold in working memory. A screen shot of this test is found in Appendix C.

Schema

Along with the three reader characteristics described above, the schema variables included two measures of schema generated by a program called Pathfinder (Schvaneveldt & Durso, 1981; Schvaneveldt et al., 1985). This program derives network structures of students' schema of a given subject. For example, if the concept of interest is stoichiometry, the program will create a representation of a student's schema of stoichiometry based on proximity data (Schvaneveldt et al., 1985). This proximity data is determined through an algorithm using relationship and similarity judgments made by the student.

The similarity judgments are made by the student using the Pathfinder's Rate program. The student is given two words in a concept pair and asked to judge on a Likert scale the degree of similarity between the two words from 0 (not related) to 9 (very related) (Schvaneveldt et al., 1985). The list of words for use in the Pathfinder program can be found in Appendix D. After the student completes a series of relatedness judgments, the Pathfinder program provides a matrix of proximity values (Schvaneveldt, 1990). These proximity values represent the nodes or concepts within the student's schema on a given topic as well as the connections between the concepts. Once this network has been established (which will be described in greater detail in Chapter 3), two possible measures based on a comparison between the student's network to a referent network can be used to evaluate the network (Gonzalvo, Canas, & Bajo, 1994). The measure used in this study was path length correlations. Path length correlations are based on the length (weight) of links between pairs of concepts. Path length correlations are found by computing the distance matrix for each Pathfinder network based on the relatedness judgments made by the student. Each of the matrices in can be reduced to a single vector. A student's vector can then be compared to an average expert's vector (referent vector) to form a ratio of shared attributes called path length correlation which provides an index of Pathfinder network similarity of 0 (low similarity) to 1 (high similarity). The second measure is a neighborhood similarity. Neighborhood similarities are a measure of the degree to which the same node in two different graphs (referent expert and participant) is surrounded by a similar set of nodes and ranges from 0 (non complimentary networks) to 1 (identical networks). The path length correlations and neighborhood similarities were used as two of the five reader characteristic variables in canonical correlation used in this study.

The three reader characteristic variables described above (logical reasoning, factual chemistry knowledge, and working memory) along with the two measures of schema (path length correlations and neighborhood similarities) as the five variables in the reader characteristic set in the canonical correlation analysis.

Measurements of Reading Comprehension Set

This study includes the following five reading comprehension variables in the reading comprehension set:

1. An encoding variable measuring the surface level of representation.
2. A lexical access variable measuring the propositional level of representation.
3. A bridging inference variable measuring the situational model.
4. An elaborative inference variable measuring the situational model.
5. A delayed summary variable measuring the situational model.

This results in five reading comprehension variables (as numbered above) involved in the reading comprehension set used in the canonical correlation. Different measurements were used for each category and will be described next.

Encoding

The encoding variable in reading comprehension studies is often measured using a student's ability to recall the text after reading a general chemistry passage and was measured in this manner during this study (Einstein, McDaniel, Owen, & Cote, 1990; Glover, Bruning, & Plake, 1982; Lorch & Lorch, 1996). Once the student had finished reading a passage he/she proceeded to a typing space and was instructed to recall and type as much of the text as possible. Each student received one score based on the number of scientific words he/she recalled for each chemistry passage. This number was averaged for all of the chemistry passages read to produce one encoding score and ranged from 0 (no scientific words recalled) to 100 (all scientific words recalled). The number of words recalled was used as an indicator of the student's ability to encode the textual information at the surface level. This score was used as one of the five reading comprehension variables in the canonical correlation used in this study. The general chemistry passages used in this study can be found in Appendix E.

Lexical Access

The lexical access variable was measured using eye tracking technology. An eye tracker records the student's eye movements as an indicator of visual attention (Just & Carpenter, 1980). Eye movements can be used as an indicator of visual attention based on several theories. The first theory outlined by Hoffman and Subramaniam (1995) states that before a person's eyes can shift; his/her attention must have previously shifted. This means that if a person's eyes have shifted to a new stimulus, the person's attention must have already shifted to that stimulus. This is important to the study of lexical access because it means we can use eye movements to indicate shifts in attention. Another important theory in the use of eye movements is the eye-mind assumption (Just & Carpenter, 1980). The eye-mind assumption states that whatever the eyes are fixated on, that is what the person is processing at that given moment. The final theory important for the use of eye fixations is the immediacy assumption. This assumption states that as soon as the eye fixates on a stimulus, the person immediately starts processing that stimulus (Just & Carpenter, 1980).

These three theories, the shift of attention prior to eye movements, the eye-mind assumption, and the immediacy assumption provide the theory behind using eye fixations to investigate cognitive processes such as lexical processing.

Along with the general assumptions made when using eye fixation data, the use of this data to investigate lexical processing requires two additional assumptions. The first is that attention during reading is allocated in a serial fashion (Reichle, Pollatsek, & Rayner, 2006). This indicates that the reader's attention is given to each word one at a time. The amount of time it takes for the reader to move onto the next word is based on the time it takes to both identify the word and the time it takes to access the lexical information relevant to that word. The second assumption is that the cognitive processes involved in the signal to shift the eye from one word to the next only occur upon completion of lexical access of the first word (Reichle et al., 2006). These two assumptions allow us to use the fixation time on a word as a representation of the time it takes for the reader to complete the lexical access of the word.

In this study fixation duration on each word in the given chemistry passage was collected using eye tracking technology. The words in the text passage that were deemed scientific by the researcher (and validated by independent chemistry instructors) were investigated. The fixation duration on scientific words was used to indicate the lexical processing of these words. The lexical processing score reflected the average fixation duration for the scientific terms. The score received was the average amount of time that was necessary for the lexical processing of the scientific words in the text. This score was used as one of the reading comprehension variables in the canonical correlation in this study.

Inferences

The inference variables were measured using a multiple choice exam developed and validated by the researcher. This methodology has been used in previous studies to successfully investigate readers' inferences (Beck, McKeown, Sinatra, & Loxterman, 1991; McNamara, Kintsch, Songer, & Kintsch, 1996).The multiple choice exam tested for the student's ability to make two types of inferences, bridging and elaborative (Anderson, 2010). Bridging inferences are those that make connections with earlier parts of the text already encoded by the reader. Elaborative inferences are made when the reader adds new information from long term memory to the interpretation of the text. Bridging inferences are often made more easily then elaborative inferences (Long, Golding, & Graesser, 1992). The bridging inference multiple choice questions were created by identifying any parts of the text being read that had disconnects with the concepts presented previously in that same text (Kintsch & van Dijk, 1978; Miller & Kintsch, 1980). This occurs when a sentence refers to a concept that was previously mentioned but not connected explicitly in the text. These disconnects were identified using a model described by Kintsch (1974) which uses propositional representations of the text. Three disconnects from each passage were randomly selected from all of those identified. These were formatted into multiple choice questions.

In addition, elaborative inferences necessary for understanding each chemical passage were identified by the researcher and validated by chemistry instructors prior to use in this study. Three elaborative inferences were randomly selected for each passage and formatted into multiple choice questions. The presence of all inferences identified, both bridging and elaborative, were validated by chemistry instructors as being present in the chemistry passages. In addition, the instructors were asked to provide any additional inferences they believed necessary for the comprehension of the passage. This process resulted in three bridging and three elaborative questions for each chemistry passage. The multiple choice inference questions (bridging and elaborative) used in this study are in Appendix F.

Each participant read six chemistry passages and answered 18 bridging and 18 elaborative inference questions. The participant received a score for both types of inference questions based on the number of questions he/she answered correctly from 0 (no questions answered correctly) to 18 (all questions answered correctly). These scores represent the student's ability to make both types of inferences (Singer, 1994). The scores were used as two of the five reading comprehension variables in the canonical correlation.

General Summary of the Text

The final reading comprehension variable for use in this study is the ability to create a general summary of the text. This variable measures the student's representation of the main points involved in the text (Kintsch, 1998). This was measured during a delayed recall task that occurred approximately 24 hours after the student's initial reading of the passage. The student was given the titles of all the chemistry passages read the previous day. The student was asked to recall the main ideas of the original passage in a summary format (Guindon & Kintsch, 1984). The student's summaries were graded using a rubric developed and validated by the researcher. The student's scores on the summaries were averaged for a single summary score ranging from 0 to 15. This score is one of the variables in the complexity of their situational model representation of the chemistry passages read. It is the fifth reading comprehension variable in the canonical correlation.

The reader characteristic set measures and reading comprehension set measures were used in the canonical correlation to investigate the relationships among the variables. The results of this analysis were also used to investigate the contribution of each variable individually. These analyses were used to answer the following research questions.

Research Questions

Using the measurements described above, the research questions investigated in this study are as follows:

1) What is the nature of the relationships among schema, logical reasoning, factual chemistry knowledge, working memory capacity and reading comprehension of a text when the content is general chemistry?
2) Are the student's schema, logical reasoning, factual chemistry knowledge, or working memory capacity a good predictor of his/her ability to
    a. encode scientific words in a general chemistry text?
    b. access lexical information involved in the understanding of scientific words in a general chemistry text?
    c. make inferences necessary to understand a general chemistry text?
        i. Is there a differential effect in bridging vs. elaborative inferences?
    d. recall the general ideas presented within a general chemistry text approximately 24 hours after initial reading?

Research Hypotheses

Based on the theories presented above and in greater detail in Chapter 2, the researcher makes the following hypotheses:

Research Question	**Research Hypothesis**
1) What is the nature of the relationships among schema, logical reasoning, factual chemistry knowledge, working memory capacity and reading comprehension of a text when the content is general chemistry?	The canonical correlation between the two sets of variables (reader characteristics, and reading comprehension) will be significant. All variables included in this study will contribute significantly to the canonical correlation.
2) a. Is the participant's schema, logical reasoning, factual chemistry knowledge, or working memory capacity a good predictor of his/her ability to encode scientific words in a general chemistry text?	The two schema variables will account for the greatest amount of variance in the student's ability to encode the textual information.
2) b. Is the participant's schema, logical reasoning, factual chemistry knowledge, or working memory capacity a good predictor of his/her ability to access lexical information involved in the understanding of scientific words in a general chemistry text?	The two schema variables will account for the greatest amount of variance in the student's ability to access lexical information.
2) c. Is the participant's schema, logical reasoning, factual chemistry knowledge, or working memory capacity a good predictor of his/her ability to make inferences (bridging and elaborative) necessary to understand a general chemistry text?	The two schema variables will account for the greatest amount of variance in the student's ability to make inferences. The two schema variables will account for a greater amount of variance in the elaborative inference variable then the bridging inference variable.
2) d. Is the participant's schema, logical reasoning, factual chemistry knowledge, or working memory capacity a good predictor of his/her ability to recall the general ideas presented within a general chemistry text approximately 24 hours after the initial reading?	The two schema variables will account for the greatest amount of variance in the student's ability to recall the general ideas.

Chapter 2 will present an overview of the literature of the theories used in this study. This includes a discussion of scientific literacy, schema, reading comprehension, and the levels of representation readers utilize.

CHAPTER TWO

> "The claim to know some scientific statement is a claim to know the process or likely process through which the statement was conceived, the degree of certainty that the field attaches to the statement, the role in reasoning the statement plays in connection with other scientific statements, and the implications of the statement's being true. If such interrelationships are missed in the reading, then the point of science is missed. The main source of both the substantive content of science and of the interrelationships within it is accurate interpretation of science text."
> (Norris and Phillips, 2003, p. 237)

In this passage, Norris and Phillips (2003) describe the importance of scientific literacy in the learning of science. In their explanation the authors convey that the learning of science cannot be thought of as rote memorization of relevant facts, but must take into consideration the role of a literacy component that describes the relationships among these facts. This component is important in the doing, teaching, and learning of science (Cromley & Snyder-Hogan, 2010). To teach students how to be a scientist or study a scientific field, we must emphasize two aspects of scientific literacy, namely, a *derived* sense of literacy and a *fundamental* sense of literacy (Ford, Yore, & Anthony, 1997; Hand, Prain, & Yore, 2001). The first aspect, the *derived* sense, is the knowledge about main facts or unifying concepts/themes of science. These concepts are specific to the different fields of study, such as chemistry, biology or physics. The second, less emphasized aspect, the *fundamental* sense is a literacy component that stresses cognitive abilities such as critical thinking, communication of ideas, and evaluation of the validity of information (Norris & Phillips, 2003). Unfortunately, many educators focus on only the scientific content, the *derived* sense, and forget the *fundamental* sense which is needed for true understanding of the material (Pressley & Wharton-McDonald, 1997b). This is a detriment to the students who are attempting to learn science as well as to the teacher who is trying to teach science effectively. No matter how much of the substantive content of science is learned, a failure to learn how to read scientific text will ultimately result in a failure to understand science (Cromley, 2009). For this reason, it is important that scientific literacy and its components be studied so that we may better understand the complex cognitive processes involved in this type of literacy.

For complete understanding of a complex process like scientific literacy, an explanation of the cognitive components involved is necessary which should include the general theories of cognitive architecture and reading comprehension. The first step in this understanding is a description of the overarching cognitive architecture involved. The cognitive architecture provides a description of how cognitive processes occur in general (Anderson, 2007). These processes can then be applied to more specific tasks such as reading. Knowledge of cognitive architecture also includes understanding the theory of schema, and how a person stores knowledge for later use. By first understanding the overarching cognitive structures, we will be better equipped to investigate and describe the processes of reading comprehension.

Cognitive Architecture

The theory of cognitive architecture is the underlying cognitive theory that drives this research project. One of the best-known descriptions of cognitive architecture comes from Brooks (1962). He described cognitive architecture through an analogy to building architecture. There are three main concepts that must be discussed to fully understand the concept of cognitive architecture, namely: *architecture, structure, and function*.

Architecture: Anderson (1983) initially described architecture as the basic operations of cognitive structure. Pylyshyn (1984) further described it as the biological substrates that achieve the basic operations of the mind. In general architecture can be thought of as the connections between the physical structure of the brain and the intended functions of the mind. The 'architect' will not necessarily live in or build the structure (which will be described next), but the building must still function as intended. Similarly, when we are discussing cognitive architecture, we are more interested in how the structures of the brain achieve the functions of the mind.

Structure: In a building, the structure can be described as the physical components of the building, for example, the wood, nails, and beams. A cognitive structure is composed of the biological components of the brain such as neurons and the connections between these neurons (Anderson, 2007). It refers both to what components are involved and where they are located.

Function: The function of the building is to enable habitation. Similarly, the function of the brain is to enable cognition (Anderson, 2007). One difference between the two is that the function of the building may involve another party, the dweller or inhabitant. In cognitive architecture the structure itself is the agent and there will be no other party involved. One measure of the effectiveness of the function (cognition) is the resulting behavior. If the behavior is desirable or effective, then the function (cognition) is effective.

Anderson (2007) analogy is illustrated in Figure 1 show the interactions of the different components in ACT-R.

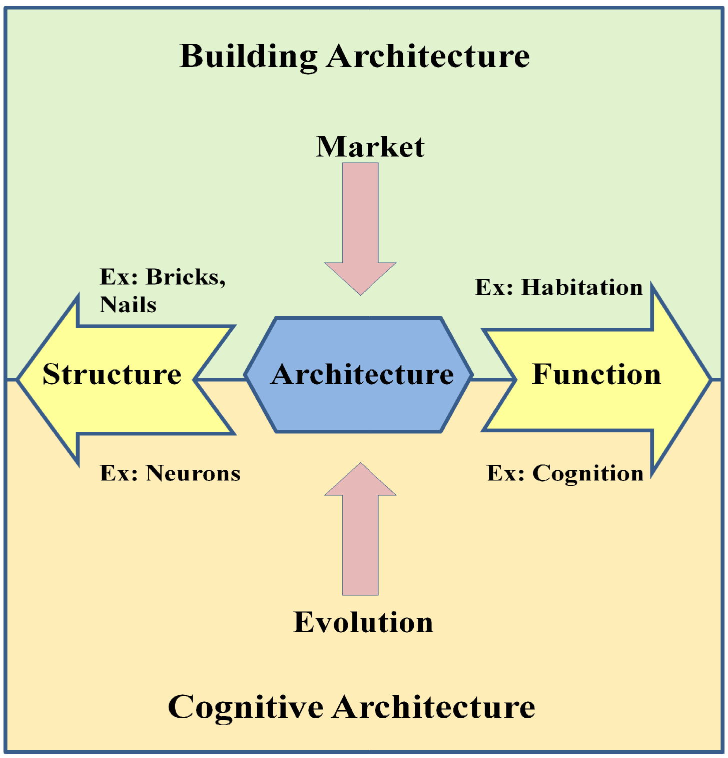


Figure 1. An illustration of the analogy between building architecture and cognitive architecture. Based on *How Can the Human Mind Occur in the Physical Universe* by J.R. Anderson, 2007, p. 6. Copyright 2007 by Oxford University Press.

This figure shows the analogy of cognitive architecture to the architecture of a building. In building architecture the market forces can have influence over the system as a whole. For instance, the people in the market who are buying homes will change how the homes are built (two vs. four bedrooms). Similarly, evolution will change the cognitive architecture of the brain. However, evolution forces take a much longer time to have an effect on the system. For these changes to take place, the resulting benefits must outweigh the cost of changing the system. Taking all these pieces into account the following is a definition of cognitive architecture, “Cognitive architecture is a specification of the structure of the brain at a level of abstraction that explains how it achieves the function of the mind” (Anderson, 2007, p.7). In this definition the phrase “specification of the structure” refers to a description of the biological components of the brain (ex: neurons). The phrase “at a level of abstraction” refers to the symbolic and sub-symbolic components of the mind that are used to abstractly describe the processes occurring in the brain.

ACT-R (Adaptive Control of Thought – Rational) is a cognitive architecture proposed by Anderson (2005). The theory can be described as a physical reality (through the structure) or as an abstraction using symbolic and sub-symbolic descriptions. Anderson (1983) describes symbols as the abstract characterization of how the brain encodes knowledge. He refers to Newell’s (1990) description of symbols as providing distal access to those knowledge-bearing structures located physically elsewhere within the brain. Our brains store related knowledge in close proximity to one another so that it can be accessed quickly and efficiently. If this organization weren’t present, the cognitive demand of even the simplest tasks would be overwhelming. Most processing is done locally, involving parts of the brain that are physically close together. However, we often need knowledge stored elsewhere and must have access to that knowledge. Symbols provide this distal access.

The sub-symbolic structure determines which symbols will be retrieved and how quickly they will be retrieved (Anderson, 2005). Each symbol is assigned a sub-symbolic quantity or value that determines in what situations it will be retrieved and how quickly this retrieval will occur. This process of providing objects with real quantities allows the system to access the most relevant and useful information (Hull, 1952).

The rational or "R" portion of ACT-R designates the system as having optimal processing in a given situation (Anderson et al., 2004). The knowledge retrieved or the actions taken will be those that provide the best reaction to the given situation. This is achieved over time by the 'tuning' of the human mind. This is not to say that the person 'logically' thinks through the possible answers and chooses the best one. It refers to the fact that over time, the processing system has determined which responses receive the greatest and fastest reward. The system chooses these rewards to be the optimal response. This results in those responses having the greatest value associated with them. A process such as this depends on the person's prior knowledge in a given situation (Anderson, 2005). The optimality will not be the same for every person and results in different responses in a given situation by different people.

The ACT-R cognitive architecture is organized in modules. Anderson (2007) has identified eight modules though he does not claim that these are the only modules in the cognitive system. The modules include: perceptual modules (visual and auditory), response modules (manual and vocal), a goal module, an imaginal module, a declarative module, and a central procedural module. These modules can be explained through the example of reading a paragraph. The perceptual modules perceive incoming stimuli from our environment, in this case, the text passage. The response modules initiate responses in the form of actions, which could be the continuation or cessation of reading. The goal module holds the current goal state of the person, for instance reading the next word or sentence (this goal is often broken down into sub goals that are represented in the goal state). The imaginal module keeps a current representation of the task at hand. The declarative module provides knowledge that has been encountered previously and which is necessary for reading the passage. This could include different types of facts. The central procedural module recognizes patterns in the other modules and responds accordingly with appropriate cognitive processes. Every module plays a role in completing a task, no matter how difficult the task is.

Human cognition occurs through the interaction of the modules described above (Anderson, 2007). Within each module, there is a large amount of parallel processing. For example the visual module is able to process a large amount of the visual field at the same time. Another example is the parallel processing that occurs in the declarative module. This module searches through vast amounts of information to produce the correct memory for further processing. Though there are large amounts of parallel processing in each module, there is also within module seriality (Anderson, 2007). This occurs because each module must select one piece of information to enter the between-module buffers. These are the buffers through which information passes from one module to another. The relevant module must select only one piece of information from each module to enter this buffer. In the example of the visual field, even though processing of the entire visual field is occurring (parallel processing), the visual module can only pay attention to one piece of information at a time (seriality).

There is also a large amount of between-module parallel processing. For instance, all of the modules can be processing simultaneously at any given time (Anderson, 2007). The declarative module can be searching for relative memories while the motor module is typing on a keyboard. This between-module processing is limited when it comes to exchanging information with other modules. All information must pass through the central production module. Often one module must wait for another module to provide information before it can continue processing. Table 3 provides an overview of the different types of parallel processing and seriality that are present in the ACT-R model of cognitive architecture:

Table 3

Parallel and Serial Processing in ACT-R

Within module parallel processing	Large amounts of information can be processed in a module at the same time.
Within module seriality	Only one piece of information from the module can be selected to enter the buffer.
Between module parallel processing	All modules can process information at any given time.
Between module seriality	The modules must communicate through a central bottleneck called the central production module. Often one module must wait for another module to finish processing before it can proceed.

The concepts of modular organizations, parallel processing, seriality, and a central bottleneck are integral pieces of Anderson's ACT-R cognitive architecture. To understand how these pieces are used in different situations, for instance reading a passage, more detail concerning two of the modules should be understood. These two modules are the central procedural module and the declarative module which are key players in a complex cognitive process such as reading.

Central Procedural Module

The central bottleneck of the ACT-R cognitive architecture is the central procedural module (Anderson, 2007). All other modules communicate with one another through this module. The evidence for this bottleneck is largely based on the Psychological Refractory Period (PRP) (Pashler, 1994). This is a delay in response when participants are given two tasks to complete. In a study by Pashler (1990), participants were given two tasks and told to respond in the order the tasks were given. The second task was initiated prior to the participants having completed the first task. The interval between the onset of the second stimulus was varied by the researcher. The initiation of the second task caused a delay in the participant's ability to respond to the first task (the PRP) as the interval between the two stimuli was decreased. In Pashler's study the significant change in the participant's response was present both when the response was the same for both stimuli, manual responses ($F_{(2,34)}$= 79.8, $p < 0.001$), and when the response for each stimuli was different, vocal and manual responses ($F_{(2,28)}$= 27.0, $p < 0.001$). This delay is evidence that a central bottleneck was present in the system that didn't allow the participants to process the tasks simultaneously.

Other theories have disagreed with this assumption citing 'near perfect time sharing' as evidence for unlimited central processing (Hazeltine, Teague, & Ivry, 2002; Schumacher et al., 2001). In the study by Schumacher (2001), participants were given a tone at different frequencies and asked to respond by saying a number (one, two, or three) in response to the frequency they heard. They were also shown an object on a screen at one of three positions. The participants were asked to push a certain button depending on the location of the object. The participants were not asked to respond vocally or manually in any specified order. The researchers found that over time, participants displayed near perfect time-sharing as shown by no significant difference in the time it took the participants' to respond to the two tasks ($t(7)=0.95$, $p > 0.10$). They interpret this as evidence of unlimited central processing.

Anderson, Taatgen, and Byrne (2005) replicated this experiment using the ACT-R model and program. They found that initially the participants did not show near perfect time-sharing, and only after time were they able to display this phenomenon. Their model identified five different modules being utilized during the experiment (two perceptual, one procedural, and two response). At the beginning of the experiment the program identified three productions in the initial processing, five productions in the finger press and accessing of outcome, and two productions in the vocal response for a total of 10 productions (which are the patterns recognized or actions taken by the central procedural module). In the earlier studies by Schumacher et al. (2001), only over time (and with practice) were the participants able to collapse those productions into one production for initial processing, one production for the finger press and accessing of outcome, and one production for the vocal response for a total of three productions. The reduction of productions necessary to produce the results, as described in the ACT-R model, was the reason for the near perfect time-sharing reported in earlier studies. This phenomenon did not occur initially but only after practice. This is an indication that the processing passed through a central bottleneck and only after time was its effects mitigated through the collapse of production rules.

Byrne and Anderson (2001) also provided evidence of a central bottleneck in their study of complex cognitive processing. They found that when confronted with complex cognitive processes (multiplication and subtraction), participants took longer to do the computations at the same time than they took to do each individually. These data support the theory of a central bottleneck in that it took longer for the participants to process the information when they were asked to complete conflicting procedures then it took to complete each individually.

While acting as the central bottleneck, the main job of the procedural module is to recognize patterns in the other modules that are processing and respond with appropriate cognitive action called production rules (Anderson, 2005). Production rules are stimulus-response bonds that depend on both the current situation and the prior experiences of the problem solver. Human cognition can be described as a series of these production rules. The easiest examples to understand or describe are those production rules in the response modules, where the production rules chosen result in an observable response. Every module has specific production rules relevant to that module that can be described even if those production rules cannot be observed outwardly.

The human mind is continuously creating new production rules. Anderson (2007) proposes the following three processes for creating new production rules:

1. The process of being explicitly taught a production rule. This is what often occurs during a teacher-student interaction. The teacher demonstrates how the student should solve a problem and the student replicates the teacher's actions.
2. Using a previous example as an analogy and applying a production rule that worked in the previous situation. Students often use this method when trying to 'map' worked examples onto new problems. Unfortunately this can be problematic when the worked example is not similar enough to the novel problem (Chi, Bassok, Lewis, Reimann, & Glaser, 1989).
3. Deducing a production rule from principles already stored in prior knowledge. This is a very difficult cognitive process. Some teachers expect their students to be able to take the principles learned in the classroom and apply them to novel situations. This can often be too difficult for students to successfully accomplish since it is very cognitively demanding.

The production rules chosen by our cognitive system are based on their effect and the time it takes to produce the effect. The reward received from using a production rule and the time it takes to receive that reward in a given situation is referred to as the production rule's utility (Anderson et al., 2004). The greater the reward a production rule receives, the greater the rule's utility and the more likely it will be selected for use in the future. This is also true for the amount of time it takes to receive the reward. The faster the reward is received, the greater the rule's utility and the more likely it will be selected for use. This results in a system that works optimally. The system selects cognitive processes that produce the greatest reward in the least amount of time.

When a production rule is first created it will have a utility of zero (Anderson, 2005). Every time it is created thereafter or used, the utility will increase and the production rule will be more likely to be selected in the future. This continues until the production rule has the highest utility in a given situation and is therefore selected over all other production rules for use.

Declarative Module

The declarative module can be thought of as a huge storage facility for information that we have encountered previously and will want to use again in the future (Anderson, 2007). It has a virtually unlimited capacity and is capable of vast parallel processing. This allows the declarative module to retrieve information very quickly (Anderson, 2005). The information stored in and retrieved from the declarative module is often thought of as factual knowledge (ex: objects, relationships, etc.). The smallest unit of declarative knowledge is defined as a proposition (Anderson & Bower, 1973; Clark, 1974; Frederiksen, 1975). A proposition is the smallest unit of knowledge that can be judged to be true. It is called a proposition because it proposes a relationship. Propositions are stored in the declarative module in chunks. The idea of chunking was first introduced by Miller (1956). He stated that only 7 ± 2 pieces of new information could be held in the short-term memory unless the smaller pieces of information are integrated into larger chunks. This allows for more storage since the chunk is treated as a single piece of information.

Chunks in ACT-R are labeled with a type and a slot. The type of chunk refers to its general category while the chunk slot refers to the exact content (Anderson, 2007).

Chunks have a few general characteristics including:

- A limited number of pieces of information can be combined into a chunk.
- A chunk has different types of slots depending on what type of chunk it is.
- Information will play different roles in chunks depending on what type of chunk it is.
- Chunks are ordered hierarchically. Smaller chunks are grouped together into larger chunks.

The large capacity, vast parallel processing, and chunking organization of the declarative module allows it to hold an almost unlimited amount of information (Anderson, 2007). The declarative pieces of information are not stored as individual production rules which would create the need for overwhelming cognitive processing. When recalled, these memories can be stated as facts, meanings, or applications. The recall of these pieces of information is flexible (Anderson, 2007). This means we can recall the information in any setting, not just the context in which we first encountered it. This flexible nature is by definition what makes a memory declarative.

As in the procedural module, declarative knowledge is selected based on the optimal system. Anderson and Schooler (1991) created an equation that represents this optimality. The equation shows that the probability of selecting a declarative memory is a function of when that memory was last used. Also, the more frequently a memory is used, the more likely it will be recalled. Both of these principles (last time used and frequency) are also true in learning (the more you use a piece of knowledge the more likely you will learn it) and retention (the more you encounter a piece of knowledge the more likely you will retain it). Context also plays a role in the selection of information. A piece of information from the declarative module is more likely to be selected if you are in the correct context for the use of that knowledge.

ACT-R and the processes involved in the functioning mind are extensive and complex as shown here in the descriptions of the procedural and declarative modules. By understanding these underlying components we are now better able discuss other organizational procedures of the mind. One of these organizational procedures that is important to the study presented here is the theory of schema. Schema theory takes the ACT-R theory of cognitive architecture and applies it to the storage and retrieval of information from the mind. It is a key component necessary to the understanding of reading comprehension. It is involved in the process of tying incoming information with information already known by the reader. Before we can understand schema in the context of reading comprehension, we must first have an understanding of the theory of schema in general.

Schema

Anderson's (2007) description of the storage of knowledge through the process of chunking is easily applied to the process through which a person's prior knowledge, ability to recall that knowledge, and storage of the knowledge affects his/her ability to utilize the information. This process has been the topic of discussion for cognitive psychologists for many years. Earlier researchers tried to describe this storage through the use of semantic networks. Quillian (1966) proposed how people stored information in categories in Figure 2 through an example of animals.

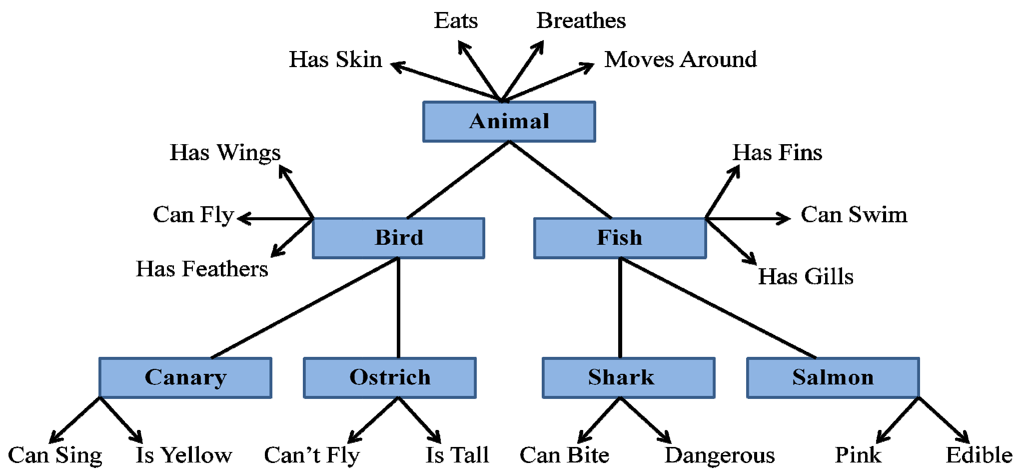


Figure 2. A hypothetical structure for a three-level hierarchy. Based on *Cognitive Psychology and its* Implications, by Anderson, 2010, p. 132. Copyright 2010 by Worth Publishers.

In this example the overarching node is 'Animal' and the 'Bird' and 'Fish' nodes fall below. Bird and Fish are thought of as subcategories of the overarching node and are thus linked to it. Each node has certain features it possesses and thus any subcategory below it should possess those features as well. Cognitive psychologists working under this theory believe that our knowledge structures are complex series of networks linked in various ways.

Unfortunately our general knowledge is not sufficiently explained by this complex series of networks, nodes, and links. For instance consider our general knowledge about dogs. This knowledge may include the following facts:

- They are a type of animal.
- They are a type of mammal.

- They can be many different breeds.
- They typically have fur.
- They are typically bigger then a breadbox but smaller then a pony.
- They can be tame or wild.
- They can work as assistants to the blind.

The type of network described above would have difficulty capturing or representing all of the nuances involved in these concepts (Anderson, 2010). For instance, the fact that dogs typically have fur does not reflect knowledge that the fur can look very different (long, short, smooth, curly, etc.). The problems with a network representation of knowledge have led researchers to propose another method for capturing this structure that seems more effective then networks. This newer method of representing knowledge is called schema (Bartlett, 1932). Through his theory of schema, Bartlett addresses the question of what processes people utilize to remember and use knowledge.

Bartlett (1932) described his theory of schema as the act of comprehending new material that requires an “effort after meaning”. Bartlett stated that the act of acquiring new information requires the person to assimilate new material into his/her existing concepts. The outcome of the process does not result in a duplication of the new information, but instead a new product in which the person’s schema and the incoming information are combined into something that is meaningful to the person. The person changes the new information to fit his/her existing concepts, or changes his/her existing concepts to accommodate the new information. While these changes occur, details of the original information are lost and the knowledge becomes more coherent to the individual.

To understand the theory of schema, we must also understand a few underlying points about schema as described by Mayer (1983). First, the concept of schema is general. It may be used in a wide variety of situations as a framework for understanding incoming information. Second, schema is a description of knowledge. It exists in memory as something that a person knows. Third, schema has structure. A schema is organized around a theme or concept. Finally, a schema contains "slots" that are filled by specific pieces of information. These four concepts, when taken as a whole, are the main components in the theory of schema. Understanding these points in general allows us to apply the schema to the storage and retrieval of information in different examples and/or situations.

The general knowledge of schema also involves the process of grouping information together categorically or chunking. The description of a schema is much like the description presented in the ACT-R framework of chunks (Anderson, 2007). Much like ACT-R's chunking, schemas also represent the categorical knowledge according to a slot framework. These slots have values and various attributes that other members of the category possess. For example, for the schema representation of a dog we would find the following (slots are given in italics with different values filled in):

Dog:

Is a: animal

Parts: Tail, four legs, body, snout

Function: Pet, hunter, Seeing Eye dog

Shape: Four legs attached to a body, which is attached to a head and tail

Size: Bigger then a bread box but smaller then a pony

The examples written here are not necessarily the only values that can fill these slots. For instance we can imagine a dog that does not have a tail or a statue of a dog that is bigger than a pony. These values are the default values we use to represent a typical dog, but they can be overwritten in certain situations if necessary.

Schemas also have a higher order hierarchical structure that influences how the information is stored (Mayer, 1983). Through this hierarchical structure, assumptions about the different parts subsumed under the higher order categories can be made. For instance we can assume that a specific dog has paws since we know that most dogs have legs and these legs have paws. Schemas are thus able to store categorizations as abstractions made from specific inferences of the concepts they represent. Therefore, if we are talking about a dog, we can use the schema to infer conclusions about what the dog will look like or act like. This abstraction also allows us to deal with exceptions (Anderson, 2010). Abstraction of the general category allows us to accept a dog without a tail as still being a dog.

This description so far has been about the storage of knowledge in general. A general description is possible because the storage of knowledge in complex schemas is much the same regardless of the information stored or the context in which it is encountered. If we understand the concepts involved in schema theory, we can then apply these concepts to individual and specific situations that can vary. The schema theory can thus be used to describe the storage of information in many different processes, for example reading comprehension.

Schema and Reading

The theory of schema described above can be applied directly to the processes of reading and comprehension of text. Chomsky (2006) describes a theory of language that applies the idea of schema to the understanding and storage of textual information. His theory is based on three basic ideas:

The distinction between surface structure and deep structure: The surface structure of a sentence is the way it is written while the deep structure is the way it is represented in memory.

Transformation rules: Language consists of a set of rules for converting surface structure into deep structure (comprehension) and deep structure into surface structure (recall and communication).

Universal grammar: Some general characteristics are shared by all language users.

This theory makes the important distinction between surface structure and memory structure. It describes how sentences are stored in the mind and how this storage is not necessarily the same as the text (surface structure). The surface structure of a text may lead to many different deep structures depending on the reader's interpretation. Mayer (1983) provides the following example. The sentence "They are eating apples" can be understood differently if "they" is taken to refer to humans or if it is taken to refer to the apples themselves. The fact that a specific surface level can lead to various different deep levels implies that a study of surface structure may not completely explain a reader's comprehension of that text.

A reader will use his/her schema in learning and remembering new information. It could therefore be beneficial to emphasize the use of schema to readers when they are confronted with new information. It has been shown that readers given parts of a schema readily infer the missing components (Bower, Black, & Turner, 1979). In fact readers often believe that they actually read what they inferred. In one experiment in this study by Bower et al. (1979), passages were read by 45 participants. The passages used in this experiment had a range in the number of cues they provided the participant. These cues prompted the participant to make inferences while reading the text. The participants were next given a recognition test in which they were asked to rate different parts of the story from one (very sure I did not read the sentence) to seven (very sure I did read this sentence). Bower found that when the original passages contained a large number of cues to make inferences, the participants were significantly more likely to "fill in the gaps" by making inferences ($F_{(2, 84)}$= 31.3, $p < 0.001$). This indicates that providing cues to the reader in the text (much like cues the reader's schema can provide) can help guide them to make appropriate inferences.

Aid in providing an overall structure for the reader can be given in various ways. This can include help with surface structure comprehension of different textual components or assistance in the deep structure representations that take the reader's prior knowledge into consideration. Some studies have shown that the simple addition of a title to the text has a significant effect on the comprehension of a passage (Bransford & Johnson, 1972). Other studies have shown that a reader's prior factual knowledge of the concept being discussed in a passage will affect his/her ability to comprehend and interpret a passage (Bransford & Johnson, 1972; Pichert & Anderson, 1977; Schallert, 1976). These studies indicate that a reader's schema will have a significant effect on the various levels of reading comprehension. This effect should be apparent in the reader's mental representations of the text and his/her ability to understand the text.

A person's schema and how it is constructed and organized directly affects the way in which he/she reads and comprehends a text. This is also true for the components and processes involved in the person's cognitive architecture including, but not limited to, the person's procedural and declarative modules. Understanding the way these cognitive processes occur in the human mind in general gives us the basis for understanding how they are applied in various cognitive processes like reading and reading comprehension. The next step in this discussion is to apply the theories described above to reading comprehension and use them to describe what is occurring in the mind during reading comprehension.

Reading Comprehension

Reading comprehension can be described in two major stages (Just & Carpenter, 1987). The first stage involves the encoding of textual information. The second stage is the process by which the words are combined to form a mental representation of the text. This representation contains the word meanings obtained by the reader as well as the relationships among the different words in the text including various inferences made by the reader while reading the text. A third stage, though it will not be investigated in this study, is the utilization stage where the reader is able to use the information in the text. One example of utilization would be use of the text to solve a given problem. The utilization of the text is an important component that should be investigated in future studies. However, we must understand the cognitive processes involved in the first two stages before we can determine how they affect performance on the third. The first stage, encoding, is the first step in reading comprehension that must occur for the reader to perceive and understand the incoming information.

Encoding

Feature analysis.

The encoding of textual information can be described using a model called feature analysis. This is a model where a stimulus is thought of as a combination of features (Gibson, 1969). For instance, the capital letter A can be thought of as two slanted lines connected with a horizontal line. This model of encoding was developed as an improvement to the original theory of template matching models (Anderson, 2010). The template matching theory was unable to fully describe the processes a reader uses to recognize various letters and words. In the template matching model, a stimulus is compared to a known template to determine if it matches and can therefore be categorized as being the same as the template. One problem with this model is that it doesn't account for our ability to encode information that does not exactly match the template. The feature analysis model addresses some of the problems encountered when using a template matching model because it is based on much smaller units of comparison. By using these smaller features we are able to encode information that may not exactly match the template. For example, all of the following A's can be categorized as such even though they don't necessarily match our default template for the typical letter A.

Another benefit of using a feature analysis model is that we are able to specify those features that are critically important to the pattern. Some features, though they are included in the pattern recognition, are not as important as other features. Feature analysis describes the way in which we are able to recognize different letters while reading. Once we are able to recognize various letters, we can then combine these letters with our declarative knowledge about each letter and determine what words are present in the text. This is the first step in encoding textual information.

Immediacy of interpretation.

One important principle in the process of encoding is the immediacy of interpretation. This principle states that the reader tries to extract meaning from each word as the reader arrives at the word (Just & Carpenter, 1987). The reader does not wait until the end of a sentence or passage to determine how he/she will interpret the word. This is evident in an experiment by Thibadeau, Just, and Carpenter (1982) where they showed that the time spent fixating on a word is proportional to the amount of information provided by the word. To do this the researchers gave fifteen passages to fourteen readers. Each passage contained an array of words including novel words (ex: flywheel) or cognitively difficult words (ex: thermoluminescence). When the researchers looked at how long the readers' spent on each word, they found that when the readers encountered novel or cognitively difficult words, they spend an average of 686 extra milliseconds beyond what would have been predicted by the word's length and infrequency. It was determined that when a reader comes across difficult or new words, he/she spends more time fixating on that word. The length of fixation in this case reflects the cognitive processes involved in understanding the meaning of the word (lexical access) rather than the length or frequency of the word.

The encoding of textual information also includes the processing of syntactic structure and semantic information (Kintsch, 1998) Syntactic structure information can come from two sources, word order and inflectional structure. In English, the dominant syntactic cue is word order though people will use both word order and inflectional structure to help interpret the meaning of a sentence. The reader must also use the meaning of the words themselves, or the semantic information, to understand text. A string of words can provide meaning to the reader even if the words are syntactically out of order based on the semantic information (Kintsch, 1998). People attempt to utilize both semantic and syntactic cues together to arrive at the best interpretation of the sentence.

When reading a text, the reader must first translate the printed text into a mental concept that he/she is able to recognize as a word that has meaning to himself/herself (Just & Carpenter, 1987). This involves two processes. The first is the process of encoding the textual information which has been described above. The second is accessing the word's meaning. These two processes often occur closely together in time. They can be distinguished from one another in situations where one occurs without the other. One example of this is the tip-of-the-tongue phenomenon. This occurs when a person has accessed the word meaning but is unable to think of the corresponding word (Just & Carpenter, 1987). Accessing the meaning of words is the second step in the reader comprehending what is written in the text. This process can be described through lexical access.

Lexical Access

Semantic memory and the lexicon.

After a reader has successfully recognized the letters and the words in the text, he/she must access the meanings of those words and create a mental representation (Kintsch, 1998). Representation of meaning of text is closely related with a reader's *semantic memory* or personal knowledge. A reader's semantic memory of a concept is unique to his/her own experiences which influence the representation of that concept in his/her mind. An important subset of semantic memory is the lexicon (Kintsch, 1974). Word concepts are entered into a person's lexicon as abstract entities that represent the concept through words or phrases. These word concepts are often denoted by being written in capital letters. This is to differentiate between the word and the word concept. For instance the word piano and the word concept PIANO are very different from one another in the context presented here (Kintsch, 1974). The word concept PIANO represents all different types of pianos as well as the meaning of the word and appropriate usage of the word. Lexical descriptions may contain sensory information (ex: HOT) as well as how words are related to one another (ex: JUMP and TRAMPOLINE). Lexical entries are defined primarily in their relationships to other entries. The relationship between the lexicon and a person's schema is very strong. The lexicon provides word meanings and word uses grouped in the schema for the person to access when creating his/her mental representation of a concept.

Understanding a text involves understanding the word concepts stored in the lexicon and relating it to the reader's prior knowledge of the concept (Kintsch, 1974). Thus to understand a text, the reader must assimilate the text with his/her prior knowledge and experiences. In the text, word concepts are joined together to create *propositions*. A proposition is a combination of word concepts (lexical items), with one word concept being the predicator and the others serving as arguments. This is much like Anderson's (Anderson, 2007) propositions in that the propositions are the smallest units of knowledge within the text that can be judged to be true. The semantic base of a text consists of ordered lists of propositions. Each proposition contains one predicator and *n* arguments. The ultimate task of the lexicon is to determine which combinations of word concepts are appropriate and therefore permissible in the situation. The reader's lexicon accomplishes this by accessing the meaning and relationships of the word concepts with one another and choosing the relation or combination that is most appropriate or useful in a given situation (Kintsch, 1998). Consider the following sentences and propositions. The propositions are written in parenthesis with the predicator written first and word concepts separated by commas.

(1) *John sleeps.*	(SLEEP, JOHN)
(2) *Mary bakes a cake.*	(BAKE, MARY, CAKE)
(3) *The man is sick.*	(SICK, MAN)
(4) *If Mary trusts John she is a fool.*	(IF, (TRUST, MARY, JOHN), (FOOL, MARY))

In (1) and (2) the predicator is a verb, while in (3) the predicator is an adjective. Sentence (4) shows the use of a conjunction as a predicator and also illustrates how proposition can contain other propositions as arguments.

Many different sentences can be represented by the same propositions (Kintsch, 1974). Which sentence structure is chosen to be included in the text is decided by the author of the sentence. For instance in the proposition (BAKE, MARY CAKE), the proposition can represent all of the following sentences:

Mary bakes a cake.

Mary is baking a cake.

A cake is being baked by Mary.

Mary's baking of a cake.

No matter which sentence is chosen by the author, the same proposition is the basis for the word concepts and therefore the lexical access. Just as the word concepts are combined to create propositions, propositions are combined to create the text base (Kintsch, 1998). The text base includes all of the information necessary for the author to get the intended idea across to his/her readers. This includes the use of relevant propositions in such a way that connections between those propositions and thus the word concepts involved in those propositions are created. The layout of the propositions chosen by the author leads the reader to make appropriate inferences necessary to encode and understand the information included in the text.

Accessing the lexicon.

Merely having a mental lexicon does not automatically supply the reader with the word meanings (Just & Carpenter, 1987). The meanings of the various words must be accessed in the lexicon and be made available to other comprehension processes. The reader must be able to access one word meaning among the many thousands of other word meanings in the lexicon. Two methods for this access have been proposed, a *search* method and a *direct access* method (Just & Carpenter, 1987). A *search* model is one where the stored word meanings are searched one at a time until the correct word meaning is located. In this model, no indexing or organization scheme is in place. This method is akin to locating a book in a library where the books are shelved randomly. Clearly, this type of search would be extremely time intensive as well as cognitively heavy. It is more likely that the accessing of meaning occurs via *direct access*.

Just and Carpenter (1987) describe a direct access model to obtain word meanings from the lexicon in which the reader has an indexing scheme that allows the reader to access any lexical entry directly. If you consider the analogy used above of a book in a library, a direct access model would be the use of an indexing system at that library. Someone familiar with that system can find the book quickly. The way in which a direct access system works is based on the *rational* component of the ACT-R cognitive architecture (Anderson, 2007). As the word is encountered, the meaning associated with the word in the person's schema is activated and given a greater utility value then other word meanings associated with that word. This activation allows for the meaning to be easily accessed and utilized by the person. Additionally, the more times a word is encountered, the faster and easier the word will be activated.

Words that are used more frequently in language are processed faster and more accurately than words that are used infrequently (Just & Carpenter, 1987). This phenomenon is referred to as the *word-frequency effect*. This effect is evident in various studies of reader's ability to recognize words during brief presentations and the time it takes to classify those words (Glanzer & Ehrenreich, 1979; Howes & Solomon, 1951). The word-frequency effect applies to the word meaning and accessing of the lexicon, not the text of the word itself. For example, a word like *bill* should be accessed more rapidly when it is applied to its more frequently used meaning (a payment) than its less frequent usage (a bird's beak). Word frequency and the other factors described above are important factors in lexical access. Each factor can affect the way a reader accesses the various word meaning involved in reading for comprehension.

As the first step in reading comprehension, encoding is important to understanding the process a reader uses to understand text. Through a description of feature analysis, immediacy of interpretation, and lexical access, we are able to understand the different components involved in the process of encoding and thus the first step in reading comprehension. This first step is the most abstract step in reading comprehension and thus the most difficult for researchers to measure and study. A problem encountered by researchers is that the various components involved in encoding, especially the accessing of word meanings, can be difficult to quantify. In many cases researchers have relied on more easily measured variables to study the complex cognitive processes involved in reading. A new advancement in this field of research is the use of eye tracking technology to investigate some of the underlying cognitive processes.

Eye movements and cognitive processes.

The processes involved in reading comprehension can be difficult to investigate as they are largely internal cognitive processes that cannot be viewed or easily measured by the researcher. One instrument that has been used more recently to investigate these processes is the eye tracker (Just & Carpenter, 1980). This instrument allows the researcher to study the eye movements of the reader. Over the past twenty years, eye movement research and eye tracking technology have provided a way for researchers to study various cognitive processes that could not otherwise be measured (Just & Carpenter, 1980). To understand how eye movements can be used as data, we must first understand how the eye works and how those workings are related to cognitive processes.

The surface of our eye is light sensitive, though it is not equally sensitive everywhere (Rayner, 1998). The visual field is about 220 degrees and that area is broken into three parts: the foveal, parafoveal, and peripheral regions. The differences between the regions lie in the sensitivity each can detect light. Objects within the visual field are only seen clearly in the foveal area. This area makes up less than eight percent of the visual field (Rayner, 1998). This is evident in the fact that while we are able to detect objects in our peripheral vision, we cannot see those objects clearly until we move our eyes to focus the object within the foveal area. The visual information in the foveal area is the information that is processed in the visual module of the cognitive architecture. The visual module maximizes the information it processes by using most of its available resources on this small area of the visual field.

To take in all of the information we need from the visual field, we must have the ability to move our eyes. These movements are necessary for three reasons as described by Rayner (1998). The first is to place the information of interest on the fovea so that it can be perceived. The second is to keep the image stationary on the fovea regardless of movement of the object or the head. The third is to prevent stationary objects from fading perceptually. These processes maximize the amount and quality of information that can be processed in the visual module and thus processed by the cognitive architecture.

When we read a text we continually make eye movements called saccades (Rayner, 1998). Saccades are rapid eye movements during which the sensitivity to visual input is reduced. Between the saccades are fixations during which our eyes are mostly still and the quality of the visual input is increased. Some characteristics of a stimulus can be identified in the parafoveal or peripheral fields of vision but to get the best and most complete information about the stimulus we must move our eyes so that we can fixate the stimulus in the foveal region (Rayner, 1998). During saccades the visual acuity is suppressed and thus we mostly perceive the world through our mind linking together the series of fixations. By moving the eye we can shift the visual input to the stimulus we are attending in the other cognitive modules.

To use eye tracking to investigate cognitive processes we must first make two assumptions. The first is that prior to a shift of the eyes occurring, a shift in attention must first occur (Hoffman & Subramaniam, 1995). This means if a person's eyes move, their attention has already shifted as well. It also means that if a person's eyes are fixated on an object, we can assume their attention is on that object. The second assumption is the eye mind connection (Just & Carpenter, 1980). This assumption has been investigated and found to be very strong in complex information processing tasks such as reading (Reichle, Pollatsek, & Rayner, 2006). Evidence for the eye mind hypothesis can be seen in studies that investigate the time spent looking at a word and how this is influenced by the characteristics of the word such as word length, frequency, and difficulty as described previously (Just & Carpenter, 1980). It is this connection that allows us to draw conclusions about encoding by studying a person's eye movements.

A second part of the eye mind connection is the immediacy assumption. This assumption states that the mind starts processing an object as soon as the eye fixates on that object (Just & Carpenter, 1980). The three assumptions presented here, namely, eye mind connection, immediacy, and attention allow us to use fixation data to investigate cognitive processes occurring during complex tasks such as reading.

Eye movements and reading.

While reading text, quick saccades are used by the reader to bring a new region of text into the foveal region so that it can be processed (Rayner, 1998). A majority of the words in a text are fixated on during reading; however, not every word is brought into the foveal field for processing. Many words are skipped and processed in the parafoveal or peripheral region. In a study by Just and Carpenter (1980) the researchers gave 15 short passages to 14 college students. Each of the passages was approximately 135 words long and was selected from *Newsweek* and *Time* magazines. The passages were shown to participants using a computer with a remote camera that was able to monitor eye fixations. Through this methodology the researchers were able to determine that content words are fixated about 85% of the time whereas function words (such as the, and, etc.) are fixated about 35% of the time (Just & Carpenter, 1980; Just & Carpenter, 1987). This study also showed that an average of 30 extra milliseconds was added for each additional letter in the word. They also found that those words with complex meanings had longer fixations then would be predicted by their length. This provides evidence that during the time the word is fixated, the reader is accessing the appropriate cognitive representations for that word. This process has also been described in the E-Z Reader model by Reichle, Pollatsek, and Rayner (2006).

The E-Z Reader model has two core assumptions (Reichle, Pollatsek, & Rayner, 2006). The first is that attention is allocated in a strictly serial fashion. This means attention is paid to only one word at a time. This occurs because the attention is closely linked to the lexical processing of the word. The time it takes to process a word while reading is not equal to the time it takes to visually input the word as a stimulus. The processing time also includes the time it takes for word identification. This requires that the reader's attention is focused on the word that is being processed. The second assumption is that processes involved in the encoding of the attended word are the signals for the initiation of the saccadic shift to the next word and for a shift of covert attention (Reichle et al., 2006). This means the signal for a shift to the next word is completion of lexical access. The fixation time on a word is therefore a representation of the time it takes for the reader to complete the lexical access of the word.

In the E-Z reader model, the lexical processing of a word is assumed to begin as soon as attention is allocated to the visual features of that word based on the immediacy assumption described by Just and Carpenter (1980). The model states that lexical processing is completed in two stages—an early stage that works as a familiarity check and a later stage in which the reader completes lexical access of the word. The stages differ because different types of lexical information about the word can become available to the reader at different points in time. For instance, the familiarity check can be based on word-form information alone while the completion of lexical access is based on meaning. The familiarity check can be thought of as the rapid recognition of the word. The completion of lexical access involves the accessing of specific information about the word, for example its meaning or uses. The fixation duration on a word is a function of the time it takes to complete these two stages. Therefore, a reader's fixation durations can be used to examine the differences in the lexical processing of texts.

Through the use of eye tracking technology, the lexical access component of reading comprehension can be investigated (Just & Carpenter, 1980; Rayner, 1998; Reichle, Pollatsek, & Rayner, 2006). The lexical access eye tracking data can be used together with other encoding data to more accurately investigate the initial steps in reading comprehension. Without these processes, the reader cannot progress to further stages in reading comprehension and therefore, cannot understand the passage he/she is reading. If the reader can successfully encode the text through feature analysis and determine the meaning and uses of the text through lexical access, the reader can then start the second process involved in reading comprehension, i.e., the development of the mental representation of the text.

Mental Representations

The formation of a mental representation of the text is an integral part of normal text comprehension (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983). This process is an automatic process that occurs in the reader's mind naturally during reading. The creation of mental representations involves the reader recognizing cues in the text such as the topic sentence, general ideas in the text, and using these to make different types of inferences necessary for the understanding of the text. These inferences are necessary to make connections both to other words in the text and to the reader's prior knowledge of the concept being presented in the text passage. Making inferences activates the reader's long term memory and schema to bring relevant information into the reader's working memory space for processing. To understand the process of making inferences, we must first understand the theory of working memory.

Working Memory and Comprehension

The theory of working memory is an expansion of initial theories of short term memory (Anderson, 2010). The initial theories of short term memory for incoming stimuli presented an extremely limited workspace in which only a small amount of incoming information could be processed at any given time (Atkinson & Shiffrin, 1968). The content capacity of short term memory is 7 +/- 2 (Miller, 1956). So how is it that a person is able to remember more than seven words from a given text? The fact is that many complex cognitive tasks exceed the memory capacities we observe in the typical cognitive laboratory experiment. The short term memory being so limited creates a need for another theory that can account for the many processes involved during comprehension. The answer to this problem comes through a theory of memory that includes a more complex model called working memory.

Long-term memory is defined as everything a person knows and remembers and is largely thought of as declarative and procedural knowledge (Ericsson & Kintsch, 1995). Just because a person knows something and that information has been stored it in his/her long-term memory does not mean the person will be able to call upon that relevant information in a given cognitive process at a given time. The information cannot affect the processing unless it is retrieved and utilized. This utilization occurs in the person's central processor called working memory (Kintsch, 1998). To affect a cognitive process, items from long-term memory must be activated for use and brought into the working memory space. The working memory is thus an active platform for processing information from long term memory during a given cognitive process. It takes into account the limited capacity of the short term memory but also allows for additional processing through retrieval of information from the long-term memory. In this way working memory provides the basis for the complex cognitive processes involved in reading comprehension.

Van Dijk and Kintsch (1983) described the various components involved in the working memory during text comprehension. These included sensory features, linguistic expressions, propositional structures, and situational models that control structures, goals, lexical knowledge, and frames involved in the text. All of these play a role in text comprehension, so they must be available to the working memory during comprehension. These processes have been shown in a study by Ericsson and Kintsch (1995) in which they found the short-term memory cues (incoming stimuli) could prompt retrieval operations from the long term memory. These retrieval operations required between 300 to 400 ms. The two main components involved in working memory according to this study were the incoming stimuli stored in the short-term memory and the information retrieved from the long-term memory (Sternberg, 1969). Working memory space allows for the incoming stimuli to be processed with the person's previous knowledge. The working memory combines the stimuli with long-term knowledge to allow for processing that in reading a text involves the making of inferences and allows for the creation of mental representations.

Inferences

After a reader has finished accessing the meaning of a sentence, the reader will utilize the words in the sentence to create a mental representation of the text (Just & Carpenter, 1987). The reader almost never passively reads a sentence by just recording the meaning. Some sentences call the reader to action through their structure alone. For instance, a question or imperative involves the reader taking some action in response. Even declarative sentences are rarely simply registered by the reader but instead require making inferences and connections. By making these inferences and connections, the reader is able to construct more elaborate cognitive representations of the text and therefore, better able to remember the textual information.

Making inferences is an important process in reading comprehension. The two types of inferences that are typically discussed are bridging inferences (or backward inferences) and elaborative inferences (or forward inferences) (Anderson, 2010). Bridging inferences make connections with earlier parts of the text already encoded by the reader. Elaborative inferences add new information to the interpretation of text. A study by Singer (1994) illustrates the differences between the two types of inferences. In Singer's study, participants were given the following three statements:

1. Direct statement: The dentist pulled the tooth painlessly. The patient liked the method.

2. Bridging inference: The tooth was pulled painlessly. The dentist used a new method.

3. Elaborative inference: The tooth was pulled painlessly. The patient liked the new method.

Participants were then asked whether a dentist pulled the tooth. This is stated explicitly in the first statement but is inferred in the second two statements. The inference in the second statement is a bridging inference where the reader must connect the dentist in the second sentence with the tooth pulling in the first. In the third statement the reader must make an elaborative inference that a dentist pulled the tooth since a dentist is not mentioned in either of the sentences in the statement. The authors of this study found that participants were equally likely to identify *A dentist pulled the tooth* in statements one and two. This means that it was fairly easy for the readers to make a bridging inference. The participants were much slower at verifying the statement when it was an elaborative inference as in statement 3. This study indicated that both inferences were successfully made by the participants, but that the elaborative inferences presented more difficulty to the participants then did bridging inferences.

The reason for the difficulty with elaborative inferences may be due to the fact there are an unlimited number of elaborative inferences that could be retrieved from the relevant schema in their long-term memory (Just & Carpenter, 1987). Several studies have been conducted that have tried to determine exactly what elaborative inferences will be made by readers. For example, one study by Long, Golding, and Graesser (1992) determined whether or not a reader would be more likely to infer information after reading a sentence that prompted the reader to do so. They found that after reading a sentence which prompted them to an elaborative inference, the reader was much faster at completing a lexical decision task related to that inference. This task involved the participants identifying whether a string of letters was a word or a nonword. If the string of letters was a word, the participant would push YES. If the string of letters was not a word, the participant would push NO. The researchers found that if the string of letters was a word that was relevant to the elaborative inference the participant would select YES significantly faster than if the string of letters was a word that was not relevant ($F_{(1,32)}=19.07$, $p < 0.05$). The researchers concluded that while reading a text that includes prompts or leads the reader to make elaborative inferences, the reader will more easily make inferences about the concept involved in the text. This means the reader is more likely to draw conclusions about the concept relevant to the elaborative inferences.

Bridging inferences are more often made automatically and thus faster than elaborative inferences (Long, Golding, & Graesser, 1992). This is most likely because bridging inferences are often critically necessary to find meaning in the text and understand the sentence while elaborative inferences are not. Individual differences in reader's reading ability can affect the reader's ability to make inferences. To make these inferences, readers must be sufficiently engaged with the text. They also must have a reading ability that is high enough to effectively encode the text. In one study by Murray and Burke (Murray & Burke, 2003) the researchers had high, moderate, and low reading ability participants read passages like the following:

> Jennie's parents had made all the arrangements for her wedding to Bob. The ceremony was to be held under the old oak tree on the lawn in front of their house. A local restaurant owner had agreed to provide the food and drink for the wedding party, and enough picnic tables had been rented to seat the guests in her parent's backyard. Everything would be wonderful if the weather would cooperate.

The students were then given one of the following two endings:

Experimental: On Saturday, the day before the wedding, the weather was threatening and the forecast was grim. Sunday morning, Jennie woke and groaned as she looked outside.

Control: On Saturday, the day before the wedding, the weather was nice and the forecast was good. Sunday morning, Jennie woke and groaned as she looked outside.

After reading one of the two passages, experimental or control, the participants were presented with a probing word like "RAIN" which is related to an elaborative inference that would only be made if the reader read the experimental condition. The participants in the experimental condition were asked to read the word. The researchers found using a one tailed t-test that those participants who were high ability readers read the word faster in the experimental condition then those participants who were low ability readers ($t_1(21)=2.52$, $p < 0.025$, $SE_{diff} = 4.39$). This indicated that the high ability readers were more easily able to make the elaborative inference than the low ability readers.

Making inferences and recognizing when they are appropriate is important to a reader's ability to create accurate mental representations of a given text. This is the second step involved in reading comprehension and is important to a reader's understanding of the text. The first two steps in reading comprehension, encoding the textual information and creating an accurate mental representation, are necessary steps in reading comprehension and must be successfully completed before a third step, utilization of the text, can occur. If the first two steps do not occur, the reader will not comprehend the text or at the very least have misunderstandings about the text he/she is reading. If we can understand how the encoding of text occurs and how the reader makes mental representations of the text then we will have a better understanding of problems that may occur during these processes. The first two steps in reading comprehension can be studied through an understanding of the various stages a reader will progress in his/her comprehension of the text. While going through the two stages involved in reading comprehension, the reader will have different levels of representation of the text. By studying these levels of representation, we can gain a better understanding of the reader's comprehension process.

Levels of Representation of Text

Successful reading comprehension results in multiple levels of representation of text in the reader's mind (Kintsch, 1998). By having multiple levels of representation, a reader increases the likelihood that he/she will be able to utilize the text appropriately in working memory or, if necessary, commit the text to long term memory. Kintsch (1998) describes a reader's internal representation of text as having three levels. The first level is the surface level of representation. This is the exact text and sentences presented to the reader. At this level the reader is not summarizing the text in any way but just representing the text exactly as it is written. The second level is called the propositional level. At this level the reader is representing and remembering connections between the words but not the actual text itself. The final level is a situational model that contains the major points of the story or paragraph. The reader has to have these three levels of representation to understand the information presented. When asked to recall the information, readers may not be able to recall all three levels. In fact, over time the surface level and propositional level may be lost and the reader will be left with the situational model in which he/she has only an overview of what the information was about (Kintsch, 1998). This higher level situational model is more durable in the reader's memory then the surface or proposition level. Success at all three levels is necessary for the reader to understand and utilize the text presented. A reader's representation of the text must be investigated at all three levels to have a complete picture of the reader's comprehension process.

Surface Level

The surface level representation described by Kintsch (1998) is based on the propositional representation of text described in his earlier work which distinguished between the microstructure and macrostructure of a text (van Dijk & Kintsch, 1983). This level of representation made by the reader is very similar to the method of propositional encoding described in the reading comprehension section of this chapter. During the act of reading, the reader will be deriving the propositions from the textbase. As described earlier, a proposition is a combination of word concepts (lexical items), with one word concept being the predicator and the others serving as arguments from the text.

The meaning of a sentence is first represented in the reader's mind by a series of words or word concepts (Kintsch, 1998). These will later be used to create propositions that include a predicate along with one or more arguments. Only those words that are attended to and perceived properly can be used to create the propositional representation (Kintsch, 1998). The predicate is a relational term and is usually a verb, adjective, or adverb in the English language. A reader's mental representation of a predicate is characterized by a predicate *frame*. The type of frame used, much like Anderson's chunks (Anderson, 2007), will specify which arguments are associated with the predicate in the reader's representation. For example, a verb frame may specify that a particular verb must have an agent, an object, and an optional goal. The categories within the frame may also have restrictions, for instance that the agent in the verb from may need to be a human. The reader's mind uses the frame to guide which pieces fit into each category. This process ensures that the reader accurately interprets the text by inserting appropriate word concepts into the category with the frame in his/her mental representation. The words or word concepts present in the text must be inserted correctly in the reader's frame for the reader to understand the intended meaning of the text.

The surface level of representation is important for the reader's understanding of the text since the appropriate use of words in different categories within the frames results in the understanding of the meaning of the sentence. At this level the reader is determining the relationships between word concepts in the text and making sense of these relationships. This level of representation is not an effective level for the reader to store the textual information in their memory (Kintsch, 1998). If this were the case, the reader would be storing all of the information in the text as the original textbase, including the exact words used. The reading of even a simple passage would present an immense strain on the reader's cognitive capacity and this strain would grow exponentially if the passage being read was complex or lengthy. The next level of representation, the propositional level, starts to decrease the cognitive load for the reader by representing the text as more general ideas rather than exact text.

Textual Proposition Level

Once the reader has successfully created a surface level of representation of the word concepts and determined the relationships and propositional frames used in the sentences, the reader moves to developing a propositional representation of the text (Kintsch, 1998). At this level, the reader is storing information as whole textual propositions rather than the words and sentences in the text. These textual propositions are much like the propositions described in the cognitive architecture section of this chapter. The textual propositions can be described as more general ideas of the text and are derived from the textbase. The meanings of entire sentences can be represented in the reader's mind as complex textual propositions. These propositions include the general idea intended by the author. Thus, textual proposition representation captures the relationships among the words and can therefore be thought of as a larger unit.

The text propositional level can also include propositions from the reader's previous knowledge that are relevant to the information in the text as inferences. These prior knowledge propositions were described in the discussion of Anderson's (2007) cognitive architecture as part of the declarative module. The integration of the textual propositions and the declarative knowledge as inferences results in the beginning of a macrostructure of the text (van Dijk, 1980). The macrostructure includes the general ideas of the text in a hierarchically ordered set. This order is sometimes directly signaled in the text, but can also be inferred by the reader. The macrostructure or propositional representation can be thought of as the ideal summary of the text. It will include the major propositions needed to understand the meaning of the text. The macrostructure is ordered hierarchically in a way that makes sense to the reader at various levels of generality. The next level of representation, the situational model, will take the propositional level and integrate a greater amount of previous knowledge contributed by the reader to create a more complete understanding of the text.

Situational Model

The final level of representation is the situational model. This includes the text-derived propositions and the propositions contributed from the reader's declarative knowledge (Kitsch, 1998). This level of representation can be thought of as the major points of the text. If the propositional representation is a summary of the text, the situational model is an outline. This level of representation is based on the overarching structure of the text. For instance, the situational model may include the presence of a topic sentence as the main point of the paragraph. This point is recognized as important and thus weighted more heavily in the reader's mental representation. Thus, the situational model's major points are ordered hierarchically to weight those pieces of information that are more important, more heavily than others.

The situation model level of representation has been identified using studies that investigate participant's ability to engage in upper level processing. A study by Guindon and Kintsch (1984) shows that all readers engage in this level of processing. In their study the researchers used a recognition priming process to investigate the situational model processing of readers. A text was given to participants that started with an initial topic sentence. For example, a text describing the training of decathloners would start with a topic sentence "A decathloner develops a *well-rounded athletic* body" and contain as one of the sentences in the body of the paragraph the statement "A decathloner also builds *up strong* hands". The participants were then presented with a target word and preceding word pair. In what the authors called macropairs, the pair corresponded to the topic sentence (ex: *develop, body*). In micropairs, both words came from the detail sentence (ex: *build, hand*). In control pairs, words that appeared in the text but in different sentences succeeded each other. The participants were asked to judge whether or not the preceding word was from the same sentence as the target word, and whether or not it was from the same proposition. The time it took the participant to make this judgment was measured. In all the examples given above the correct response to the different relatedness judgments was yes (affirmative). The researchers found that the priming effect was substantially larger for the macro words than for the micro words. This indicates that the participants were automatically weighting the macrostructures of the situational model components of the text, such as topic sentences greater than the propositional level components. Well-organized situational models are crucial for understanding and remembering text. In later studies researchers investigated the creation of situational models further by evaluating participants' abilities to create situational models when the consistency of the text was not coherent (didn't follow a logical storyline or thought process).

Texts that are locally coherent but contain global inconsistencies are read more slowly and remembered less well (Albrecht & O'Brien, 1991; Albrecht, O'Brien, Mason, & Myers, 1995). In a study by Albrecht and O'Brien (1991), participants were given passages to read. Each passage contained three main concepts. The participants were then given a 'probe' sentence to read. These sentences were either found in the original text and considered true or they were not found in the original text and considered false. The participants were asked to identify each probe as true or false (in the original text or not). The researchers found that those probes that were coherent to and consistent with the text were identified more quickly than the probes which presented inconsistencies with the original text ($F_{(2,66)}$=10.67, $p < 0.05$). This result indicates that coherency in the macroscopic structure of the text is important to the participant's ability to read and comprehend the text. Though the first two levels of representation are important to the understanding of the text, the situational model is what allows the reader to synthesize the textual information with his/her previous knowledge to create a representation of the general ideas that is meaningful.

Kintsch (1998) provides a good overview of the three levels of representation, surface, propositional, and situational model, through the description of a reader reading a passage about Argentina and Brazil. Suppose a text describes the comparison between the two countries in terms of their geography, agriculture, industry, and population with each comparison written as a separate paragraph. Also suppose the reader already knows a lot of information about Argentina but has no knowledge nor cares to know much about Brazil. The levels of representation in the reader's mind might consist of the following:

1. Surface level: The reader has a representation of the words and the relationships between those words. This reflects the order in which the sentences were presented.
2. Propositional level: The reader has a propositional level derived from the text. This level includes the four major subdivisions corresponding to the four major topics discussed in the text, population, economy, geography, and culture. Sublevels below each of these general ideas including summaries of the differences between the two countries represented. The general ideas of the sublevels would be hierarchically lower than the overviews of the four major topics yielding an overall structure for the storage of the information.

3. Situational model: Because the reader knows nothing about Brazil and does not care to know about it, the situational model for this section would be very similar to the propositional level. The reader would not have additional information to add to the situational model. In the reader's situational model of the discussion involving Argentina, he/she might add elaborations or have the ability to make inferences. This allows the reader to reorganize his/her mental representation of the information involving Argentina more thoroughly than his/her representation of Brazil.

Through this example, we can see the different levels of representation that a reader might construct. Each level plays an important role in the reader's understanding of the passage. Without one of the levels, the reader loses some of the information necessary for understanding. The reader's ability or lack of ability to create the three levels of representation will directly impact his/her ability to understand and utilize the textual information.

All three levels of representation (surface, propositional, and situational) are necessary for the reader to successfully understand the text presented (Kintsch, 1998). Only through successful representation at all three levels will the reader comprehend the text. This is not to say that successful representations will always lead to comprehension. Without one or all of the three levels (surface, propositional, and situational), the first two stages in reading comprehension (encoding and creating a meaningful mental representation) cannot occur. The creation of the three levels of representation involves the reader's ability to encode the incoming stimuli (text). It also involves the reader's ability to retrieve information from his/her current schema of the content from long term memory and use this to make inferences. Finally it involves the reader's ability create a mental outline of the main topics of the text and integrate this with his/her prior knowledge. All the components described thus far are vital to the reader's understanding of the information being presented.

The processes involved in reading comprehension described above are general enough to be applied to the reading of many different types of texts with different formats and different types of content. Therefore, these general theories can be applied to the reading of scientific texts in the learning of science. The reading component of learning science is a major component and should be well understood by educators in this field. The reading of scientific text and scientific literacy are skills needed by all scientists and by extension, to students learning to become scientists. The present study will investigate the processes involved in the reading and comprehension of scientific texts in an effort to understand the students' scientific literacy.

Science Literacy

Over the past half decade the role of literacy in the teaching and learning of science has gone through several changes. In the 1960s the US educational system underwent reforms that nearly eliminated text from elementary science instruction (Yore & Treagust, 2006). This resulted in a curriculum where no textbooks or supplementary information books were used and student learning centered solely on science inquiry models. This method produced students who understood the basic concepts in the scientific fields but were largely unable to read, write, or discuss scientific concepts. In the last two decades the attitude towards the role of language in science has changed and more language instruction has been included in the various teaching methodologies (Norris & Phillips, 2003). This process has occurred slowly as it can be difficult to convince current and future science teachers that language is an essential part of science learning and should be taught to their students.

In the English language, the concept of literacy is described through two major premises (Norris & Phillips, 2003). The first more basic description of literacy is the ability to read and write. The second meaning of literacy is based on a person's knowledge, learning, and education. These two definitions are related. In some cases it is possible to have one without the other. For instance, a person can be knowledgeable without being able to read and write. A person can be very skilled at a trade through trial and error or apprenticeship and never have read a text on the subject. This can be true of a single case; however, when we think of a disciplined body of knowledge as a whole, it is difficult to ignore the connection between the two types of literacy. Text is used to record, describe, and explain ideas to other people in the field. Without text this transmission of information would be greatly hampered and therefore, the progress of a field would be hampered. If a person cannot read scientific text, he/she may be severely limited in the depth of scientific knowledge that he/she can acquire.

The argument for the two types of science literacy above has been described in detail by Norris and Phillips (2003). They define the two types of literacy as the *fundamental* sense of scientific literacy and the *derived* sense of scientific literacy. The *fundamental* sense of literacy is the ability to read and write when the content is science while the *derived* sense of literacy is being knowledgeable, learned, and educated in science. Studies in the past have often paid greater attention to the *derived* sense rather than to the *fundamental* sense and have thus omitted an important part of scientific literacy.

In previous studies scientific literacy has been described in some of the following ways: (a) knowledge of the substantive content of science and the ability to distinguish science from nonscience (Mayer, 1997; Shortland, 1988); (b) understanding science and its applications (DeBoer, 2000; Hurd, 1998; Shen, 1975; Shortland, 1988); (c) ability to use scientific knowledge in problem solving (American Association for the Advancement of Science (AAAS), 1993; DeBoer, 2000; Hanrahan, 1999; National Research Council, 1996; Norman, 1998); (d) understanding the nature of science including its relationships with culture (DeBoer, 2000; Hanrahan, 1999; Korpan, Bisanz, Bisanz, & Henderson, 1997; Norman, 1998). Notice that these theories provide detailed descriptions of many aspects of the *derived* sense of scientific literacy, but few mention any points that could be described as the *fundamental* sense. In other words these theories focus almost exclusively on knowledge of concepts in science in order to be considered scientifically literate but they do not discuss the ability to effectively read and write text containing these concepts. Anderson (1999), described the focus of many science educators as being entirely on the hands-on experience while neglecting the importance of reading and writing in science. He stated that reading and writing are the mechanisms through which scientists create, share, and negotiate the meanings of their work.

None of the previously cited work on scientific literacy investigates the *fundamental* sense of literacy and its importance to the learning of science. Norris and Phillips (2003) describe *fundamental* scientific literacy as having more than a functional relationship with respect to science. In other words, the reading of text is not only a tool for the storage of science but also a fundamental process in the learning of science. The reading and writing of text is an essential part of science without which a person cannot be truly engaged in the scientific field or the learning of science.

Fundamental Sense of Literacy

In the reforms of the 1950's, the reading of scientific text took on the perception of being a passive activity (Yore, Craig, & Maquire, 1998). This is why educational practices moved towards eliminating the reading of scientific text and concentrated on a more "hands-on" approach. This practice was not often challenged by educators because for many whose primary area of work is not reading, the process of reading seemed too easy. To those who find reading easy and are knowledgeable in the content area (as many teachers of science are) the reading of a scientific text seems like a simple task. However, a study by Pressley and Wharton-McDonald (1997a) found that this is not necessarily true for science novices. They found that simply because a student can decode the words within a text does not necessarily mean he/she comprehends that text. The decoding of the text is only the first step in a series of processes that must occur for the comprehension of the text. Many students considered to be "good" readers in the traditional sense may not fully comprehend the science text given to them (Haas & Flower, 1988). That is, they may be able to read and know the words, identify and locate information, and recall content, but be unable to analyze, summarize, or critique the text when asked to do so.

In many classrooms, educators may foster the student's inaccurate belief that understanding the scientific text is the ability to say or read the words correctly (Haas &

Flower, 1988). This occurs when the instructor requires that the students be able to repeat information verbatim to succeed but do not ensure that the student understands the information he/she is repeating. Unfortunately for the students, skilled word recognition is not sufficient in the complex field of science. Decoding is not sufficient for comprehension and thus not sufficient for the learning and/or utilization of scientific text. This idea has been reiterated often through research (Anderson, 1985; Goodman, 1985). These studies report that reading is not merely knowing the individual terms and vocabulary. Determining the meanings of words does not yield the meaning of the propositions, and determining the meanings of propositions does not yield the meaning of the entire text.

Inferring meaning from a text involves the integration of the text information with the reader's schema or previous knowledge (Norris & Phillips, 2003). The integration involves both the text and the reader's previous knowledge being integrated in a new way that results in the interpretation of the text. Not all interpretations of a text made by the reader will be the same, nor will they all be of similar quality. More than one good interpretation of a single text can exist, but no text can be written so explicitly that interpretation by the reader is not necessary. Interpretation of the text is necessary regardless of what is being read, though the purpose of reading may vary across text types and reading contexts.

This theory of reading implies that there is some type of relationship among the author, the reader, and the text that connects them (Kintsch, 2004). The reader makes many judgments about the text including what the intention of the author was in writing the text, what the text implies, and what the value is of what is said in the text (deCastell, Luke, & MacLennan, 1986). Olson (1994) further described literate thought as the conscious representation and deliberate manipulation of thinking and reading. While the reader reads the text he/she is making assumptions and inferences. Literate thought is characterized by the reader's recognition of those processes. By doing this the reader is actively constructing his/her mental representation of the text rather than passively piecing together the definitions of the words.

Reading comprehension depends on the background knowledge of the reader, as well as the decisions made by the reader regarding the information's relevance (Cain & Oakhill, 2004; Norris & Phillips, 1994; Spearritt, 1972). It requires active construction by the reader of new meanings, contextualization, and inferring of authorial intentions (Craig & Yore, 1996; Yore, Craig, & Macquire, 1998). For this theory of reading to be valid, reading must be understood to have a number of different features (Norris & Phillips, 1987). First, reading proceeds through a number of stages with each providing a more concise interpretation. These stages consist of steps though the steps are not necessarily followed in a specific order. Reading is also interactive. The reader actively negotiates among his/her mental representation of what is being said, the textual information provided, and his/her background knowledge. Finally, reading involves a continuous self check to determine if the interpretations made by the reader are complete and consistent. Through this process the reader is able to determine which interpretations are most likely to be correct and/or useful.

Scientific text, as it is described here, and the ability to read that text is one of the key aspects of teaching and learning science (Norris & Phillips, 2003). This should lead educators to teach not only the scientific content present in the text (*derived* sense), but also how to accurately read the text for deeper understanding (*fundamental* sense). To teach this process, we must first understand what is occurring in the reader's mind during the two initial steps involved in reading comprehension outlined in this chapter, encoding the text and creating a mental representation of the text (Just & Carpenter, 1987; Kintsch, 1998). These steps provide the foundation for comprehension of any type of text, including scientific text. Thus, an in-depth understanding of the processes involved in encoding and creation of mental representations will provide for better understanding of reading comprehension, a major component in scientific literacy.

As described earlier in this chapter, the processes involved in reading comprehension of scientific texts can be investigated through the measurement of a reader's representation of scientific text (Kintsch, 1998). The student progresses through three stages of mental representation of the text namely, the surface level, the propositional level, and the situational model. The progression through each of these levels is necessary for understanding, though not all three levels will be retained by the user for future use. The three levels of mental representation can be used as measurable indicators of the underlying cognitive processes involved in reading comprehension when the content is scientific in nature. The complexity of the surface level can be used as an indicator of the student's ability to encode textual information. The student's ability to complete the tasks involved in the creation of the propositional level, namely lexical processing, can indicate the beginning of mental representations involved in reading comprehension. Finally, the student's situational model can be used to indicate his/her success in making inferences and creating a summary of the text that can be used in the future. By investigating the levels of representation and how they are affected by various reader characteristics such as schema, logical reasoning, factual knowledge, and working memory, we can better understand the processes involved in reading comprehension when the text is scientific in nature. This may lead to a better understanding of how we can teach students to become scientifically literate.

Outline of Study

This study is designed to investigate the effect of a student's reader characteristics, namely schema, logical reasoning, and factual chemistry knowledge on reading comprehension when the text presented is scientific in nature, specifically chemistry. Reading comprehension was investigated through the measurement of the students' levels of textual representation. This includes the measurement of four processes namely, encoding, lexical accessing, inference making, and creation of situational models. These processes were used as indicators of students' reading comprehension.

Reader Characteristic Variables

Schema.

As stated above, schemas are the way in which a person stores knowledge about a subject for later retrieval and use (Bartlett, 1932). The extent of knowledge a person has of a subject and the relationships between concepts within that subject affect how complete and structurally complex his/her schemas will be (Chi, 2006). Someone with extensive knowledge on a subject (an expert) will have very different schemas then someone without that extensive knowledge. Expert's schemas will contain more concepts in them and also more connections and relationships between those concepts (Goldsmith, Johnson, & Acton, 1991). To capture the underlying organization of the person's schema, a methodology using relation and similarity judgments has been used by the cognitive psychology community

(Schvaneveldt & Durso, 1981; Schvaneveldt et al., 1985; Schvaneveldt, 1990). Peoples' judgments of the relatedness between the members of pairs of concepts within a subject have been shown to capture the underlying organization of their schemas of that subject (Gonzalvo et al., 1994). This procedure produces a matrix of proximity values. These values represent the degree of relationships between a pair of concepts called nodes.

A weight associated with the strength of the relationship between two nodes is associated with each link (Gonzalvo, Canas, & Bajo, 1994). This weight reflects the distance between the nodes. For example, consider the subject of mammals. Nodes like lion and bear may be located closer to one another then lion and cat since their habitat are very different from one another. Nodes can be linked together directly (lion—bear) or through an indirect path (lion—cat—mouse).

In this study the students' schema of general chemistry concepts was investigated using a Pathfinder program (Schvaneveldt, 1990). This program searches through the nodes to determine the closest direct path between nodes (Gonzalvo, Canas, & Bajo, 1994). A link remains in the network only if it is involved in a minimum length path between two nodes. Through this program, Pathfinder finds the latent structure combining the nodes (concepts). The Pathfinder program has been shown to capture categorical relations (Schvaneveldt & Durso, 1981) and representational shifts in experts and novices (Goldsmith, Johnson, & Acton, 1991b; Schvaneveldt et al., 1985). The Pathfinder structure allows for two separate measures to be calculated (Gonzalvo et al., 1994). These include the correlation among students' and experts' number and weight of links between pairs of concepts (nodes), and the degree to which the same node in two graphs (for example expert and novice) is surrounded by a similar set of nodes.

In the present study, the students' schemas were analyzed using the two measurements described above. First, a similarity score based on correlation between the number and weight of links between concept pairs for each participant and a referent expert network was calculated. Second, the degree to which a node in the participant's structure is surrounded by similar nodes as compared to a referent expert network was evaluated. The use of the Pathfinder program in this research provides an objective technique for the representation of participants' schemas and the assessment of those schemas for completeness and structure. The use of this program will be discussed in detail in Chapter 3.

Logical reasoning, factual chemistry knowledge, and working memory.

Not all differences in a students' ability to successfully encode text, access lexical information, make inferences, and create situational models can be attributed to their schemas. Other variables have also been shown to affect the processes involved in reading comprehension. These include logical reasoning, factual knowledge, and working memory capacity (Bransford & Johnson, 1972; Cain & Oakhill, 2004; Ericsson & Kintsch, 1995; Kintsch, 1994; Palladino, Cornoldi, De Beni, & Pazzaglia, 2001; Pichert & Anderson, 1977; van Dijk & Kintsch, 1983). These variables are also measured in this study to investigate their effect on reading comprehension. Logical reasoning, factual chemistry knowledge and working memory will be used along with the schema variables as the reader characteristic set during the statistical analysis.

Logical reasoning ability has been effectively measured in chemistry classes using the Group Assessment of Logical Reasoning (GALT) test (Bunce & Hutchinson, 1993). This test was administered to students to determine their logical reasoning ability prior to their participation in this study. Students' factual chemistry knowledge was analyzed using a multiple choice chemistry content exam developed and validated for this research. Working memory capacity is often tested using memory span tests (Anderson, 2010). In this study working memory was evaluated using a digit span test in which students are asked to repeat lists of digits. The number of digits in each list increases by one until the student can no longer accurately repeat the list. Through this method the number of items each student can hold in working memory is determined.

The reader characteristic variables described above were analyzed to determine how they affect reading comprehension. Reading comprehension was analyzed through the measurement of the various processes involved. These include the following: encoding, lexical access, inference making (bridging and elaborative), and the recall of a general summary of the text.

Reading Comprehension Variables

Encoding.

Encoding of text at the surface level is often measured using recall tasks (Einstein, McDaniel, Owen, & Cote, 1990; Glover, Bruning, & Plake, 1982; Lorch & Lorch, 1996). These studies used free recall of text to investigate encoding in various reading situations. The reading situations varied in both the way the text was presented (summary sentences, titles, etc.) and the length of the passages used. In general, recall tasks provide researchers with information about which words the reader is able to successfully encode. The words the reader encodes are affected by both the connections the word has to other words in the sentence and by the importance of the word as perceived by the reader (Fletcher & Bloom, 1988; Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975; Kozminsky, 1977). In this study the students were asked to recall a chemistry passage after they read that passage. The number of scientific words the students recalled was used as an indicator of their ability to encode the text.

Lexical access.

In this study lexical access was investigated through the use of eye tracking technology. Eye tracking analysis for the investigation of lexical processing is becoming more popular as the technology advances (Just & Carpenter, 1980). Its use is supported in this type of investigation by three theories described in detail in the Lexical Access and Reading section of this chapter. The first is the theory of attention which states when a reader is reading a text, before the eye moves to from one word to the next, the reader's attention shift's to the second word (Hoffman & Subramaniam, 1995). This means that the reader's attention has moved to the new word before the eyes move to that word. Eye movements can therefore be used to indicate attention shifts. The second theory is the eye mind assumption (Just & Carpenter, 1980). This theory states that the eye fixates on what the mind processes. This allows us to investigate the mind's processing using eye fixations. The third theory is the immediacy theory. This theory proposes that the mind immediately starts processing as soon as the eye fixates on an object (Just & Carpenter, 1980). These three theories, when taken together, support the use of eye fixations as indicators of the processing occurring in the reader's mind.

Eye tracking technology and the study of fixations is used in reading comprehension studies to investigate lexical processing. The eye fixations are used as indicators of the cognitive processing that occurs when a reader is accessing the meaning of the words written in a text. This is possible due to the assumptions outlined by Reichle, Pollatsek, and Rayner (2006) described earlier in this chapter. These assumptions include the seriality of reading and the fact that completion of the cognitive processes involved in lexical processing is the signal to the mind to shift attention to the next word. These two assumptions allow us to use the fixation length on each word as an indicator of the amount of lexical processing the reader uses to access the word's meaning. The eye fixation data can be used to indicate the length of cognitive processing involved in lexical access.

Students were eye tracked while reading chemistry passages. Their fixation data was collected using the eye tracker. The fixation durations on scientific words in the passages were measured to determine whether differences in length of fixations exist between students. This data was used to identify differences in the lexical processing of the relevant concepts by students with varying schema structure.

Inferences.

This study used multiple choice questions to investigate the students' ability to make bridging and elaborative inferences. Bridging inferences are connections with earlier parts of the text already encoded by the reader (Kintsch, 2004). An elaborative inference adds new information to the interpretation of the text. This new information is retrieved from the reader's prior knowledge (Kintsch, 2004). Using multiple choice inference questions has been used in studies to identify a reader's success in making inferences. Multiple choice questions were used in this study to investigate the students' ability to make bridging and elaborative inferences. The creation of these inference questions will be described further in Chapter 3. Students' success in answering these questions was used as an indicator of their ability to make inferences necessary for comprehension of the text.

General summary of text.

A situational model is a combination of the text-derived propositions and the reader's prior knowledge of the relevant information into a summary of the major points in the text (Kintsch, 1998). This includes a summary of the main ideas and provides the reader with a summary of the text that includes all of the main ideas but does not include the storage of the text itself. It allows the reader to store the important information in the text without overloading his/her cognitive processes with the actual text itself (van Dijk & Kintsch, 1983). The situational model is a connection of cues given by the text with information from the long term memory of the reader (Ericsson & Kintsch, 1995). These connections allow the reader to store only the important information given in the text. After the connections with long term memory have been established, it is no longer important for the reader to retain the surface and/or propositional level in their working memory. This allows the reader to store and later retrieve the information in a way that does not cognitively overload the reader's mind.

To investigate the students' situational model in this study, the students were asked to give a summary of the important information they read in the original text. These summaries were written approximately 24 hours after the original reading of the text passages occurred. This amount of time was chosen so that the surface and propositional level representations had time to decay leaving only the situational model. The summaries typed by each participant were evaluated using a rubric to determine their accuracy and completeness. The resulting score was used as an indication of the students' ability to create accurate and complete situational models of the chemistry text.

Contributions to the Field

The study presented here will contribute to research in the chemical education field by providing a better understanding of the way students' schema, logical reasoning, factual chemistry knowledge, and working memory affect their reading comprehension of scientific text. It will provide details about the students' ability to encode text, retrieve word meanings through lexical access, make inferences, and create summaries of the text. Increasing students' ability to be successful in the processes involved in scientific literacy may increase their science achievement (Cromley, 2009; O'Reilly & McNamara, 2007). While increasing students' science achievement is the ultimate goal, it is necessary to first understand the underlying processes that affect reading comprehension and thus scientific literacy. This is the purpose of the study presented here. By understanding these processes, the chemical education community may better understand where difficulties occur during the reading and comprehension of chemistry text. These difficulties can indicate where in the comprehension process students need assistance in developing their reading skills (Britton & Gulgoz, 1991; Kintsch, 1998; McNamara, Kintsch, Songer, & Kintsch, 1996). Before instruction can be devised, we must first identify the different components involved and determine if and where problems exist.

An in-depth study of the reading comprehension components involved in science literacy such as the one presented here, may identify areas where students have difficulty with the processes necessary for comprehension. Once these problem areas have been identified, chemical educators can alter their teaching practices to address the problems (Cromley & Snyder-Hogan, 2010). If chemistry educators can be made more aware of the various components involved in scientific literacy, they may be better equipped to help their students' reading comprehension of scientific text. This study increases the knowledge of those components and identifies problem areas that instructors should be made aware of. This increased awareness may result in better instructional practices to increase scientific literacy (Otero, Leon, Graesser, 2002).

The methodology proposed in this study includes new measurement tools for various components involved in scientific literacy. These include the Pathfinder program used for the study of student schema and eye tracking technology for the study of lexical processing (Gonzalvo, Canas, & Bajo, 1994; Just & Carpenter, 1980; Rayner, 1998; Reichle, Polletsek, & Rayner, 2006). The study of underlying cognitive processes can be difficult to measure as they do not result in not outwardly measurable action of the student. The use of eye tracking technology to study lexical accessing of word meanings allows the researcher to have a better understanding of the student's processing then other more subjective measurement tools (Just & Carpenter, 1980; Rayner, 1998; Reichle et al., 2006). Similarly, the use of the Pathfinder program allows the researcher to obtain a better representation of the student's schema then the student could provide himself/herself (Gonzalvo et al., 1994). This representation of schema is less affected by subjective input by the student or researcher which can occur during other schema representation measurements such as the creation of a concept map (Goldsmith, Johnson, & Acton, 1991; Schvaneveldt et al., 1985). The use of eye tracking technology and the Pathfinder program in this study can expand the use of these methodologies in the chemical education community.

The research and literature outline in this chapter support the investigation of the two research questions and subquestions involved in this study as shown below.

1) What is the nature of the relationships among schema, logical reasoning, factual chemistry knowledge, working memory capacity and reading comprehension of a text when the content is general chemistry?

2) Is the student's schema, logical reasoning, factual chemistry knowledge, or working memory capacity a good predictor of his/her ability to

   e. encode scientific words in a general chemistry text?
   f. access lexical information involved in the understanding of scientific words in a general chemistry text?
   g. make inferences necessary to understand a general chemistry text?
      ii. Is there a differential effect in bridging vs. elaborative inferences?
   h. recall the general ideas presented within a general chemistry text approximately 24 hours after initial reading?

The methodology used to collect the data for each of the variables included in these questions is discussed in Chapter 3.

CHAPTER THREE

Introduction

The purpose of this quantitative study was to use a canonical correlation to investigate the effects of students' schema, logical reasoning ability, factual chemistry knowledge, and working memory capacity on their reading comprehension when presented with a passage from a general chemistry text. Reading comprehension variables included the students' ability to encode the textual information, access relevant lexical information, make necessary inferences, and create long term summaries of the text (situational models) using the textual information given. Quantitative methods were used to investigate the variables in both sets, reader characteristic and reading comprehension. The variables were analyzed using a canonical correlation. The results, presented in Chapter 4, indicate whether the reader characteristics account for a significant amount of variance in student responses on the reading comprehension variables. These data were used to determine the types of relationships that exist between the variables. Prior to the student portion of the study, experts were tested using the Pathfinder program to 1) select the topics to be used in the chemistry passages and 2) to create referent expert networks that could be used for comparison with the student networks. The results from this expert portion was used to determine which topics would be used in the student portion and to analyze the students' results.

Expert Participants

The experts were asked to complete the Pathfinder portion in the same way the student participants would be directed later in the study. A detailed description of this process can be found below in the student participant portion of this chapter. The expert portion of the study was conducted for two purposes. The first was to determine which two of the four topics initially selected by the researcher (Atoms, Ions, and Molecules, Stoichiometry, Chemical Bonding, and Thermodynamics) would produce the best Pathfinder networks. The two topics that performed the best on the measures of interest (described in Chapter 4) were then used in the remainder of the study with the student participants. The results of this analysis indicated that Atoms, Ions, and Molecules, and Stoichiometry were the two topics most appropriate for use in the student portion of the study. The second purpose was to develop an average referent expert Pathfinder network for each of the two chemistry topics. These averaged networks were used as referent expert networks for comparison with the student-derived Pathfinder networks to determine the quality of the students' network. The methods of analysis involved in the expert study will be described in detail in Chapter 4.

Student Participants

Participants in this study were students who were attending The Catholic University of America and were currently enrolled in or had been previously enrolled in a chemistry course. All levels of chemistry courses were considered appropriate for the study from high school through senior level undergraduate chemistry courses. By using a wide range of students, a varying range of structural knowledge, factual chemistry knowledge, logical reasoning, and working memory demonstrated by the students would be evaluated.

Students were invited to participate through a short verbal presentation in either the chemistry course they were enrolled in or a chemistry organization such as the chemistry club. Students were reimbursed for their participation in both the first and second session. For those students who were eligible to obtain credit for their psychology course, participation enabled these students to receive research credit through the psychology department in place of reimbursement.

After permission was obtained from the students, background information such as SAT scores and chemistry background were released by the university (SAT) or the student (chemistry background). The scores were used to ensure that the sample was representative of students with varying levels of chemistry background (obtained from the student) and aptitude (obtained through the university). The students' critical reading SAT scores obtained through the university were evaluated as an indicator of reading ability. This was used to ensure that the students represented a range of reading abilities. An overview of the background questions/tests used to ensure a wide range of students participating in this study can be found in Table 4.

Table 4

Background Information

Background Information	**Test**
1. University demographics	1. What year in school is the student? 2. What major, if any, has the student declared? 3. What is the student's age? 4. What is the student's gender?
1. Chemistry demographics	5. Survey: Part 1 a. What chemistry courses has the student taken? b. When did the student take these courses? c. What grade did the student receive in the courses? (Appendix G)
2. Aptitude and past achievement (obtained through the university)	6. SAT score and high school GPA (The College Board, 2005)
3. Visual information	7. Survey: Part 2 a. Does the student wear glasses or contact lenses? b. Does the student have any other visual conditions they may affect the eye tracking? (Appendix G)
4. Reading ability (obtained through the university)	8. Critical reading SAT score (The College Board, 2005)

Students were solicited during various points in the fall semester until at least 40 students agreed to take part in the study. The multiple solicitations were done to control the number of students actively involved in the study at any given time since a limited number of students could do the eye tracking portion of the study each day. Having the students enter the study in groups allowed the researcher to ensure variation in the sample based on aptitude and reading ability scores described above. The final number of students with viable data was 41. This number provided enough variability in the results for comparisons to be made.

Measures and Scores for Analysis

Demographic Information

Students were asked a series of demographic questions to determine whether or not their data was appropriate for use in the analysis. This information was also used to ensure that the group chosen for the study was representative of the student population enrolled in general chemistry courses. The demographic information was collected via a paper-pencil survey and included the students' age, gender, year in school, and major. The survey also included the students' self-reported chemistry background including the chemistry courses they had taken or were taking; when they took these courses; and the grade they received in the courses (either completed or to date). The demographic information also included questions about the students' vision. The questions asked if the student was wearing corrective lenses and/or glasses and whether the student had any other visual conditions (such as thick eye glass lenses or Strabismus) which might affect participation in the eye tracking portion of the study. No students were found to have an eye or visual condition that would interfere with participation in the eye tracking portion of the study. The demographic questions used in this study can be found in Appendix G.

Aptitude

With IRB approval and student permission, students' SAT scores and high school GPA were obtained from their student records. The SAT is widely accepted as a benchmark standardized assessment of students' aptitude (Camara & Echternacht, 2000). This includes their critical reading ability, mathematical reasoning, and writing skills. The SAT is used by many universities, along with other information, to assess students' readiness for college and to predict students' future performance in their college career. A study using 151,316 high school students, found that the SAT test had a correlation of 0.53 with students' first year grade point average at their institution (Kobrin, Patterson, Shaw, Mattern, & Barbuti, 2008). This correlation was increased to 0.62 when the students' high school GPA was also included in the correlation. An additional study by Wilson (1983) found that SAT scores combined with the students' high school GPA was a better predictor of students' college performance (cumulative GPA over four years) then either score alone. This indicates that the combination of SAT score and high school GPA is a good predictor of the students' aptitude and success in college.

The current SAT and its sub-sections (critical reading, mathematical reasoning, and writing skills) have reported reliability coefficients of approximately 0.90 (College Board, 2011). The SAT score and high school GPA were used in this study as an indication of the students' aptitude and prediction of overall college performance. Possible scores on the SAT range from 600 to 2400, combining test results from three 800 point sections including math, critical reading, and writing. If a student was admitted to the university based on his/her ACT score rather than a SAT score, they refused the option to release their scores, or they were admitted without test scores, their data was not included in the collection of data that ensured a wide range of abilities (SAT score and GPA). Their data was, however, still included in the canonical correlation analysis. This was possible because the SAT and GPA scores were not included the canonical correlation and thus the student's data on the ten variables involved in the canonical correlation (variables in the reader characteristic set and the reading comprehension set) were still viable for the main portion of the study. This occurred in only three cases, a number that would not have a major affect on the distribution of scores of the SAT/GPA analysis.

Reading Aptitude

Student's reading aptitude was determined by obtaining their SAT critical reading sub scores. This section of the SAT includes reading passages and sentence completion tasks (The College Board, 2005). The passage-based reading test measures the students' comprehension of what is stated or implied by the passage. The sentence completion questions test measures the students' vocabulary and understanding of sentence structure. The subscore students received on this portion of the SAT was used as a measure of the students' reading ability (Camara & Echternacht, 2000; The College Board, 2005). The critical reading section of the SAT has a reported reliability coefficient of 0.90 (The College Board, 2005). As stated previously, this indicates a high degree of internal consistency within this sub-section of the SAT exam (Cronbach, 1951). Again, three people were not included in this analysis for the reasons described above. This missing data were determined not to have a major affect the distribution of reading aptitude scores.

Reader Characteristic Set

Logical reasoning.

To measure the students' logical reasoning ability, the Group Assessment of Logical Thinking test (GALT) was administered (Roadrangka, Yeany, & Padilla 1982). The GALT test was administered via paper and pencil and consists of twelve items which measure the

Piagetian reasoning level of the students. The test is untimed and required the students to select an answer for each question as well as a reason for their answer from a list of responses in a multiple choice format (in the first ten questions). The last two questions required the students to determine lists of combinations given a set of objects. This test is designed to measure not only the students' ability to answer correctly but also the reasoning they use. The Piagetian categories covered by the test include conservation of mass, control of variables, proportional, probabilistic, correlational, and combinatorial reasoning (Nurrenbern, 2001; Piaget, 1972). The questions range from testing at Piaget's concrete operational phase through the formal operational phase.

The GALT test has been shown to predict logical reasoning abilities as well as grades assigned by teachers in science and mathematics (Bitner, 1991). The validity and reliability of scores on the test are well established using students from grade six through college by Roadrangka, Yeany, and Padilla (1983). The test's Cronbach's alpha reliability coefficient is 0.85 (Roadrangka et al., 1983). Other studies using this test with different populations report similar reliability coefficients (Bitner, 1991; Bunce and Hutchinson, 1993). The results of the GALT test rate the students' logical reasoning ability with a score from 0-12. A high score (generally 10-12) is interpreted as the student functioning at a higher logical reasoning level. A low score (generally 6 or less) is interpreted as the student functioning at a lower logical reasoning ability level. Lower scores can indicate that the student may have difficulties understanding abstract concepts. This may mean students with a low GALT score could have trouble with advanced chemical concepts which are often abstract in nature.

The students in this study could receive a score from 0 (low logical reasoning) to 12 (high logical reasoning). The GALT test used in this study is found in Appendix A.

Factual chemistry knowledge.

Students' factual chemistry knowledge was measured using a multiple choice chemistry test with the topics used in this study (Atoms, Ions, and Molecules, and Stoichiometry). The factual chemistry knowledge test was created using publicly released Test Item Bank questions from the American Chemical Society's Exams Institute. These questions were used in previous exams in general chemistry courses in both high school and college (American Chemical Society Division of Chemical Education Examinations Institute, 1992). Questions are released for public use when the Exams Institute retires the questions and new exams are created. Twelve questions from each of the two chemistry concepts (Atoms, Ions, and Molecules, and Stoichiometry) were selected from the Test Item Bank. These items were combined to produce a 24 question multiple choice chemistry test, covering the concepts involved in this study. Twenty-four questions were compiled for this test in preparation for the potential removal of items based on validity and reliability testing with a separate set of chemistry students from those who participated in the full study.

Content validity was determined by having chemistry experts review the content and evaluate its appropriateness for each chemistry concept. The experts were given a list of the chemistry concepts and asked to determine whether the multiple choice questions adequately covered that concept. Suggestions provided by these experts were considered by the researcher who then made changes to the questions.

Item statistics using a separate group of students than the students in this study were used to determine the reliability of each item on the factual chemistry knowledge test. These statistics included a difficulty index and a discrimination index for each item. The results from these statistics were evaluated to determine which, if any, questions should be revised or removed prior to their use in the main study.

The difficulty index (p) of each item is the percentage of students who respond correctly to an item. This index was calculated for each item using the following equation:

$$p = \frac{\text{number correct}}{\text{number of answers}} \times 100$$

The lower mathematical limit of p is 0 which will occur if no student answers the item correctly. The upper limit is 100 which occurs if all the students answer the item correctly. The larger the value of p, the easier the item (American Chemical Society Division of Chemical Education Examinations Institute, 1992). As suggested by the test makers, items with p factors greater than 80 or lower than 20 were rejected for use in this study as they are either too difficult or too easy.

The discrimination index (r) measures the performance on an item for students who did well on the test overall relative to those who did poorly. This index provides a mathematical expression of the fact that on a well designed question, the top students should perform successfully more often than those students whose overall achievement is lower. The discrimination index of an item would be at its maximum if every top performing student answered the item correctly and every low performing student answered it incorrectly. The discrimination index of an item would be at its lowest value if all the low-achieving students answered a question correctly, but all of the high-achieving students answered it incorrectly. The *r* index was calculated by subtracting the number of right answers on that item given by students in the bottom 25% of the achievers from the number of right answers given by students in the top 25% of achievers. This number was then divided by the number of students in the top achieving group.

$$r = \frac{\text{high group correct} - \text{low group correct}}{\text{number in top achieving group}}$$

The higher the value of *r,* the more useful the item is at discriminating between high and low achieving students. If the *r* is either 0 or 1, the discrimination index is zero and no discrimination can take place. Items with an *r* of 0.25 or greater were accepted as having good discrimination (American Chemical Society Division of Chemical Education Examinations Institute, 1992). Those with values below 0.20 were discarded after the initial testing.

The validity testing described here was completed with 21 students early during the fall semester. Table 5 presents the results from this study by question.

Table 5

Item Statistics for Multiple Choice Chemistry Content Test

Question/Topic[a]	Difficulty Index[b]	Discrimination Index[b]
Q1 : AIM	43	0.86
Q2:S	62	0.71
Q3:S	57	0.19
Q4:AIM	57	0.71
Q5: S	48	0.86
Q6: AIM	57	0.43
Q7: AIM	90	0.14
Q8:AIM	71	0.43
Q9: S	67	0.86
Q10: AIM	43	1.00
Q11: S	52	0.43
Q12: S	71	0.29
Q13: AIM	67	0.57
Q14: AIM	86	0.29
Q15: S	38	0.57
Q16: AIM	76	0.43
Q17: S	76	0.71
Q18: S	57	1.00
Q19: AIM	67	0.43
Q20: S	43	1.00
Q21: AIM	86	0.22
Q22: S	52	0.13
Q23: S	38	0.17
Q24: AIM	67	0.86

[a] The two topics involved in this study have been abbreviated as the following: Atoms, Ions, and Molecules (AIM); and Stoichiometry (S).
[b] Values that do not fall within the acceptable range are highlighted.

As seen in the table above, five questions from the original 24 were removed from the multiple choice chemistry content test due to their performance in an unacceptable range on either item difficulty or discrimination index. This resulted in 19 questions suitable for use in the main portion of this study. Though the two topics did not have an equal number of questions (AIM=10 and S=9) the difference was small enough that the researcher chose to proceed with 19 questions. Through the processes described here the final version of the test that was used with students in the main portion of the study was determined to be statistically valid and reliable for this population. The students received a score of 0 (low factual knowledge) to 19 (high factual knowledge) on this test. The final test is included in Appendix B.

Working memory capacity.

Working memory refers to the space in which students attend to incoming information and combine it with information from long term memory needed for completing a task (Atkinson & Shiffrin, 1968; Baddeley, 1992). The capacity of this space impacts the amount of information students are able to effectively pay attention to and use. Working memory capacity was tested using a memory span test (Anderson, 2010). This type of test determines how many stimuli (often random numbers or letters) students can immediately repeat after being first presented with these stimuli. This type of test has been found to have reliability coefficients of approximately 0.90. Such tests have also been found to have similar results to other working memory tests (Ryan & Ward, 1999). In this study, students were asked to complete an electronic working memory task called a digit span test. In this test students were given a random list of digits. These digits appear one at a time on the computer screen. Once students had viewed the digits they were asked to type the list from memory in a space provided on the computer. This process was repeated with subsequent lists increasing in length by one digit. The students continued until they made two mistakes in repeating the list of digits at which point the test was concluded. The students received a score from 0 (no items held in working memory) to the greatest number of items they were able to successful hold in their working memory. This test was administered via a free online working memory website that includes this test in its database (Cognitive Fun, 2009). A screen shot of the test can be found in Appendix C.

Schema.

One of the reader characteristic variables of interest in this study is the students' structural knowledge or schema of various chemistry topics. For students' to be considered knowledgeable in a chemistry topic they must understand the interrelationships among the concepts within that topic. These relationships are learned through the incorporation of new facts into prior knowledge or the modification of schema to accommodate the new knowledge. While learning the chemistry concepts in class, students continuously modify the way their schemas represent and organize the various pieces of information. Studies have shown that the more knowledgeable a person is about a given subject, the better organized the information is within his/her schema (Chi, Feltovich, & Glaser, 1981). As described in Chapter 2, this organization or knowledge structure is an important variable in students' cognitive processes.

Educational researchers have addressed the question of how to accurately represent the organization of a student's knowledge for many years. There is no known instrument or test that can produce an exact replication of a student's structural knowledge. One attempt to determine the nature of students' knowledge structures used essays (Norman & Rumelhart, 1975). Using essays to ask students' to compare and contrast ideas can provide some information about their structural knowledge and the relationships they perceive among concepts. The use of essays, however, can be limited as other factors such as the organization and representation of knowledge in the structural knowledge can affect the way novices respond to the essay questions (Diekhoff, 1983). Using essays also requires subjective interpretation of results by the researcher which could threaten the validity of the conclusions.

Recently a new method for determining knowledge structures has been developed and tested (Schvaneveldt & Durso, 1981; Schvaneveldt et al., 1985; Schvaneveldt, 1990). This method is a graph-theoretic technique used in a computer program called Pathfinder. This program derives the network structures from proximity data. The proximity data is determined by looking at relationship and similarity judgments made by students. A student is asked to judge the relatedness or similarity between members of concept pairs (key terms). This process provides a matrix of proximity values in which each value represents the degree of a relationship between a pair of concepts. Algorithms are then used to reveal the underlying dimensions of the students' structural knowledge on which their judgments are based. The dimensions of this space represent the main properties along which the concepts within the domain are organized.

The concepts used for the relatedness judgments in this study were key terms from each chemical concept (Atoms, Ions, and Molecules, and Stoichiometry). These key terms were created using the vocabulary words in various chemistry texts (Brown, LeMay, Bursten, & Burdge, 2003; Kotz, Treichel, & Weaver, 2006; Silberberg, 2000). The vocabulary words in each text were cross referenced to create a list of terms for each concept that were common to all texts. These lists included between 25-30 key terms for each chemical concept. The lists were given to two chemistry instructors who determined whether the terms in the list adequately covered the given chemistry concept. Any key terms suggested by the instructors for addition or deletion were evaluated by the researcher and added or removed when appropriate.

The lists were presented to the students in key term pairs. The students were asked to judge the relatedness of each pair of words on a scale of 1 (completely unrelated) to 9 (completely related). The full set of instructions presented to the students can be found in Appendix H. The results from the Pathfinder program were analyzed by first converting the key term ratings to proximities (dissimilarities). This is done by subtracting each rating from the maximum possible (9). When a student gives two concepts a high relationship rating and that rating is subtracted from the maximum (9), the result is a low proximity value. If a student gives two concepts a low relationship rating and this rating is subtracted from the maximum possible (9), this results in a high proximity rating. This process results in low proximity values representing similar concepts and high-proximity values representing concepts judged to be dissimilar (Schvaneveldt, 1990).

The Pathfinder program uses the proximity values in an algorithm. This algorithm organizes data by eliminating those links that are not the minimum path between two concepts through the following procedure. The Pathfinder network is created using a q and r parameter equal to n-1 and infinity respectively (where n=the number of concepts). A path in the Pathfinder network consists of a number of nodes and connecting links. The length of the path is defined by the r parameter. This is based on the Minkowski r-metric. The length of a path defined by the r parameter is a function of the weights associated with the links in the path. This is calculated from the ratings provided by the student. As *r* decreases, links are added to the network. When the *r* is set to infinity the number of links in the network is maximally reduced. The parameter *q* defines the maximum number of links in a path and also affects network density. The parameters $r = \infty$ and $q = n-1$ generate the simplest Pathfinder network and require only ordinal assumptions to be made about the distance estimates.

Concepts or nodes in the Pathfinder network can be linked directly to one another or linked indirectly through a multi-node path. The Pathfinder algorithm searches through the nodes to find the closest direct path between nodes or concepts. A link remains in the network only if it is the most direct path between two concepts. The most direct path could be a direct node to node link, or a multi-node path. As long as it is the shortest path it remains in the network. All other links between those two concepts are removed from the network by the computer program. The transformation from proximity data (determined by the program based on students' relatedness judgments) to a Pathfinder network is illustrated below in Figure 3 with a data set taken from Schvaneveldt et al. (Schvaneveldt & Durso, 1981):

Proximity Data

	A	**B**	**C**	**D**	**E**
A	0	1	4	2	3
B	1	0	1	4	5
C	4	1	0	5	6
D	2	4	5	0	4
E	3	5	6	4	0

Pathfinder Network

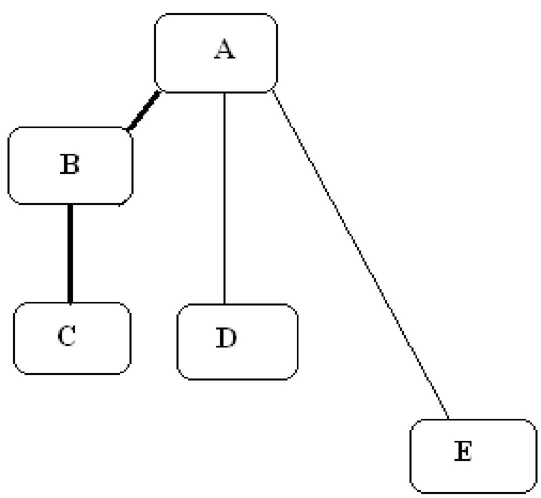


Figure 3. Transformation from proximity data to Pathfinder network using data from general semantic networks. Based on a *Paper Presented at the Annual Meeting of the Psychonomic Society* by Schvaneveldt, 1981.

In Figure 3 direct links exist between elements that are most closely related (have a proximity value of one). These links exist between A and B and also between B and C and are represented by bold lines in the figure. Links are added between the nodes A and D and also between A and E though these links are longer than the previous links since they represent weaker proximity values. Those links that represent a multi nodal path have the weakest proximity values, for example B and D or A and C. Through this process the Pathfinder is able to determine the most basic representation of the student's structural knowledge. The resulting network represents the students' knowledge of different concepts and how those concepts are linked together through relationships. One benefit of using Pathfinder is that it does not force a hierarchical solution, however if a hierarchical representation exists in the students' knowledge structure, it will be included.

The network that results from this process is then compared to the referent expert network on two measures, namely, the path length correlations and the neighborhood similarities.

In this study, a path length correlation score was computed between each student network and the average referent expert network according to the following procedure. As described above, two nodes in a single network are connected by a pathway between them which can either be close (directly linked) or far (two or more links apart in the path). The distances were calculated by computing the number of links constituting the shortest path between pairs of concepts in the network. These values were represented by a distance vector (the weight of the path). A network with n nodes may be represented by a vector of $(n^2 - n)/2$ distances. Two networks can thus be similar or different in terms of the number and weights of the connection between nodes. Path length correlation was determined in this study by calculating the correlation between the distances in two networks, student and referent experts' (Goldsmith and Davenport, 1991). The path length correlation of each student's (averaged over both topics) and the referent experts' network provided an index of network similarity with 0 (low similarity) and 1 (high similarity). High correlations represent high similarities between networks and low correlations represent low similarities between networks. Therefore a high path length correlation score would indicate that the student's knowledge structure network is similar to the referent experts' network while a low correlation would indicate that the student's knowledge structure is dissimilar to the referent experts' network.

The second assessment of the similarities between two networks used in this study is neighborhood similarity. Neighborhood similarity is measured by determining whether a set of nodes surrounding a specific node are the same set as in another network. The set of nodes that are within distance 1 (immediate neighbors) from a particular node (key term) provide information about the structural property of a network and can thus be compared to determine network similarities. This measure ranges from 0 to 1 (Goldsmith and Davenport, 1991). A student score of 0 would indicate the student's network is not similar to the referent experts' network while a score of 1 would indicate that they are identical.

The two resulting scores from using the Pathfinder program, path length correlation and neighborhood similarities, were used as two of the five reader characteristic variables in the canonical correlation used in this study. The results from this study will help determine if any differences are present in the way students are able to encode textual information in the chemistry passages, access the lexical meaning of the text, make inferences necessary to understanding the text, and recall the general ideas involved in the text after a delayed period of time. These analyses will determine if differences present in the students' encoding, lexical processing, inference making, and/or summary of the passages can be attributed to their existing schema (path length correlations and neighborhood similarities).

Reading Comprehension Set

Reading passages and chemistry concepts.

The reading passages used in this study were selected from general chemistry texts used in college level general chemistry courses (Brown, et al. 2003; Kotz, et al., 2000). Two concepts were chosen for use in the study (Atoms, Ions, and Molecules, and Stoichiometry) based on the responses of the expert chemists described previously. Three paragraphs were selected from the chemistry texts on the two concepts used in this study. The paragraphs selected can be found in Appendix E. These paragraphs contain between five and eight sentences, a short title, and a number of scientific words (those that are specific to the chemistry concept) and conversational words. These words were identified by the researcher and validated by chemistry instructors. Chemistry instructors were asked to read the passages to determine whether the researcher had correctly identified all of the scientific and conversational words in the passage. The Chemistry instructors were also asked if the chemistry passages fit with the concepts being tested. Three passages were selected for each of the two concepts for a total of six passages.

Encoding.

Encoding was measured through a recall task. The students were asked to read each of the nine chemistry passages until they felt the passage was fully understood. The amount of time needed for this process was determined by the student. Once the student had finished reading, he/she was shown a screen with the passage title and a space in which he/she could type as much of the passage as remembered verbatim. The student was then asked a series of inference questions in a multiple choice format (described in further detail below). Once each student finished the multiple choice questions he/she proceeded to the next chemistry passage. This process was repeated until each student had completed all six chemistry passages.

Students received a score based on their ability to recall the scientific words in each chemistry passage they read. The scores for all nine passages were averaged resulting in one score for the encoding analysis. The students received a score based on the percent of scientific words they were able to recall from 0 (no scientific words) to 100 (all scientific words). This score was used as a reading comprehension variable in the canonical correlation.

Lexical processing.

To test the students' lexical processing of the text, an eye tracker was used to evaluate how the students read the text. The eye tracker data was only used during the reading of the chemistry text passages, not during the relatedness judgments in the Pathfinder program or the answering of inference questions. Using the eye tracker provided information about the reading process that could not otherwise be observed by the researcher or self reported by the students. The data collected using the eye tracker provides a more objective view into the cognitive processes at play while the students are reading the chemistry text.

Eye tracking technology.

With recent advances in computer technology and programming, eye tracking has become an accessible method for collecting student data. It provides a non-invasive way to investigate some of the cognitive processes occurring in the mind that otherwise would not be easily accessible to the researcher. Eye tracking data is generally collected on student's eye fixations and the saccades that connect those fixations. A fixation is a point at which the eyes are generally immobile and fixed on one location (Rayner, 1998). Fixations are linked together by saccades which are quick jumps of the eye from area to area. During saccades vision is generally suppressed therefore, the majority of the information a student encodes and/or processes while viewing a stimulus occurs during the fixations.

Data on student's viewing patterns, including fixations and saccades, were collected using a Tobii T120 eye tracker (Tobii Technology, 2010). The hardware components of the instrument are mounted directly into a 17 inch computer monitor. By having the hardware mounted directly into a computer monitor, the eye tracking technology is nonintrusive to the student's natural reading process. The students can interact with the computer and monitor much like they would any other computer.

On the eye tracker, cameras are mounted behind a sun blocking light filter that is largely invisible to the students. The eye tracking instrumentation uses near-infrared diodes that reflect light off the student's corneas. The reflected light and other user characteristics are collected by the cameras mounted in the monitor. The system then uses a series of algorithms to analyze the data to produce a three dimensional gaze point of each eye on the monitor (Tobii Technology, 2010). In addition, the following data found in Table 6 are collected by the instrument:

Table 6

Eye Tracking Data Collected by Tobii Program

Data	Description/Units
Gaze position (for each eye)	X/Y coordinates on the monitor of each pupil
Distance from student to monitor	Millimeters
Time stamp	Milliseconds (measured from start of data collection)
Validity code	System's confidence in the recorded data (0-4 scale with 0 representing excellent data and 4 representing missing or incorrect data)

The data collection rate is 50 Hz which means that 50 gaze points were collected each second for each eye. This rate has been determined to be the most appropriate for reading analysis (Tobii Technology, 2010). Each student was seated approximately 60-75 cm from the monitor as directed by the instrument's developer. Calibration of the eye tracking instrument for the student was performed at the beginning of each eye tracking session (Tobii Technology, 2010). The calibration process involved the student viewing a series of moving dots on the monitor. These stimuli appeared in different positions and moved on the monitor during a 60 second test. Once calibrated, the eye tracking instrument is reported to be accurate to within 0.5 cm between the measured and reported gaze position of each eye (Tobii Technology, 2010). This one calibration was sufficient as long as the student did not largely reposition himself/herself during the study. An additional computer monitor was attached to the eye tracker where the researcher can view the student's gaze patterns. The researcher's view of the eye tracking process alerted the researcher if a student had repositioned himself/herself too much and the system needed to be recalibrated. The student was informed that he/she could rest between each chemistry passage if desired, but not to reposition himself/herself while resting.

Lexical access and processing.

Both the eye tracking instrumentation and protocol described above are useful in reading research because of the mind-eye connection (Reichle et al., 2006). This connection allows us to draw conclusions about the student's cognitive processes by evaluating the way his/her eyes address the text presented. In this study the eye tracking technology was used specifically to evaluate the student's lexical processing of the text. The utilization of eye movements to investigate lexical processing was described in detail in Chapter 2.

The time each student spent on a word in the text is directly related to the time it took him/her to access the meaning and relationships relevant to that word. By recording the fixation time spent on each word, the student's lexical access of the words was measured. The words in the chemistry text were broken down into a series of Areas Of Interest (AOIs). These AOIs were the size of each scientific word plus a 2 mm border.

After each student was eyetracked while reading the scientific texts, the fixation lengths for each scientific word were analyzed using the AOIs. The fixation lengths were then divided by the number of letters in the word to determine the time spent on each letter. These lengths were averaged across all passages and both topics. In this methodology, a longer length of time indicated a student having a more difficult time accessing the lexical information. To make a positive relationship between the student's fixation durations and his/her ability to access the word meanings, the averaged fixation duration per letter was subtracted from one. This meant that a larger score indicated faster lexical processing. In this methodology, each student received a fixation duration score for the time spent on scientific words. This score was used to indicate the lexical processing of students in the canonical correlation.

Inferences.

Students' ability to create inferences were measured using a multiple choice inference test that included bridging and elaborative inference questions. To determine the number of inferences necessary in each passage, Kintsch's theory of reading was utilized (Kintsch & van Dijk, 1978; Miller & Kintsch, 1980). The program developed in these studies simulates the way in which a reader encodes text and constructs a mental representation of the text. In the program the first sentence's propositions (the smallest unit that can be judged as true) are entered. Next, the second sentence's propositions are entered and the program checks for coherence between the two sentences. Coherence is found when the second sentence contains an idea (proposition) previously mentioned in the first sentence. The program then inputs the third sentence and checks for coherence, and so forth until the end of the text. Points at which coherence is *not* found between two sentences are points where the reader must make a bridging inference to link the new sentence with an idea from a previous sentence. In this manner, the bridging inferences present in the text were identified.

As stated previously, the passages were also given to chemistry instructors who were provided a list of chemistry content elaborative inferences created by the researcher. The instructors were asked to determine if the inferences were necessary to understand the passage. They were also asked to identify any elaborative inferences that were necessary but not included in the list. The elaborative inferences created by the researcher and validated by the instructors along with the bridging inferences described above were used to create a multiple choice inference test. The questions were written so that one of the answers completes the inference while the distracters do not. Not all inferences in every passage were included in the test as that would have created an extensive test which could overload the student. Three elaborative and three bridging inference questions were randomly selected from all those identified by the researcher and chemistry instructors for each passage. This resulted in a total of six questions per passage. The bridging and elaborative inferences used are included in Appendix F.

The inference test was used to evaluate students' ability to identify and make both bridging and elaborative inferences. The students received two separate scores for bridging inferences and elaborative inferences each ranging from zero (no questions answered correctly) to 9 (all questions answered correctly). These scores were used in the canonical correlation to determine whether the inference scores were affected by the reader characteristic variables.

General summary of text.

Approximately 24-48 hours after the students completed the first session's testing, they were asked to participate in the third and final session of testing. Session three provided an opportunity for each student to type a summary of the passages after being given the short title for the chemistry passages. This follow-up testing provided additional information about whether or not the students were able to effectively encode the textual information and store that information in their schema. It has been shown that information is stored long term in a more abstract level including the general ideas rather than the surface level replication of the text (Kintsch, 1998). Differences between the students' encoding and storage of information that would not necessarily be found directly after the students read the passages may be more evident after a 24 hour delay.

The students were given the title of all six passages they read the previous day and instructed to write a summary of each passage. The students were also told that this summary should include the main topic and any ideas they felt were important to the passage. The summaries created by the students were graded using a rubric created and validated by the researcher through the following process (Appendix I). Prior to use in the study, the rubric was sent to chemistry instructors. These instructors were asked to determine whether the rubric adequately evaluated the delayed summary. Any suggestions provided by the instructors were considered by the researcher and, when appropriate, changes were made to the rubric. After students had finished writing the general summary, two people (the researcher and another chemistry expert who was an instructor or working chemist) used the rubric to grade each summary. Inter-rater reliability was evaluated to ensure the graders' scoring was not significantly different from one another. The results from this analysis are described in Chapter 4. The score on each summary ranged from zero (poor summary) 15 (very good summary) and was averaged across the six passages and two chemistry topics so that each student had one summary score. The scores were used in the canonical correlation to determine if the students' delayed recall of a general summary was affected by reader characteristic variables.

Statistical Analysis

To analyze the data collected in this study a canonical correlation statistical analysis was used. As described previously, this analysis included two sets of variables, the reader characteristic set and the reading comprehension set. The reader characteristic set included two schema variables (path length correlation and neighborhood similarity), logical reasoning, factual chemistry knowledge, and working memory. The reading comprehension set included encoding, lexical access, bridging inferences, elaborative inferences, and the delayed general summary. The results from this analysis were used to evaluate relationships among the variables measured.

A canonical correlation is a multivariate generalization of the Pearson Product Moment Correlation coefficient. It is appropriate for use when the data consists of a number of continuous variables in one set and a second set of continuous variables. To understand how the canonical correlation works, assume $v=\mathbf{Yb}$ represents a linear combination of a set of p dependent variables. Let $u=\mathbf{Xa}$ represent a linear combination of a set of q predictor variables. The canonical correlation chooses the components of the vectors **a** and **b** so that the Pearson Product Moment Correlation between the two sets of variables is as large as possible. The value of the maximum correlation is the canonical correlation R_c, u and v are the canonical variates, and **a** and **b** are the canonical coefficients (weights). In other words, the canonical correlation chooses the weights to apply to each variable so that the linear combination of each the sets is as highly correlated as possible.

In this study the first set of variables is the reader characteristic set including two measures of schema (S1 and S2), logical reasoning (LR), factual chemistry knowledge (CK), and working memory (WM). The linear combination of this set can be represented using the following equation:

$v = \mathbf{a}\,S1 + \mathbf{b}\,S2 + \mathbf{c}LR + \mathbf{d}\,CK + \mathbf{e}WM + C$

where **a, b, c, d,** and **e** are the weights corresponding with each variable and C is the constant term in the linear equation.

The second set of variables is the reader comprehension set including encoding (E), lexical access (LA), bridging inferences (BI), elaborative inference (EI), and delayed general summary (DGS). The linear combination of this set can be represented using the following equation:

$u = \mathbf{f}E + \mathbf{g}LA + \mathbf{h}BI + \mathbf{i}EI + \mathbf{j}DGS + D$

Where **f, g, h, i,** and **j** are the weights for each variable and D is the constant term in the linear equation. The canonical correlation determines the weights (**a-j**) that result in the highest correlation between *v* and *u*.

Once the canonical correlation (R_c) has been calculated, it can be squared (R^2) to give the percent variance in the students' responses described by the combined set of variables. This and other results from the canonical correlation will be described in detail in Chapter 4. In this way each reader characteristic variable is evaluated to determine what percent of the variance in responses it accounts for in the reading comprehension variables. Separate regressions can also be evaluated to further investigate each variable in the reader comprehension set and determine what percent of the variance in that variable is due to the different reader characteristic variables.

To review the variables presented in the Measures and Scores for Analysis section of this chapter, Table 7 provides the following: an overview of the variables involved in the study; the variable sets in the canonical correlation; the instruments used to measure the variables; and the variable names selected by the researcher which will be used in the next section of this chapter as well as Chapter 4 and 5.

Table 7

Statistical Analysis Used

	Statistical procedure used	**Variable groups**	**Variables involved**	**Variable Name**	**Instrument used**
Testing for range of students	Normal distribution	N/A	- aptitude - reading ability	Not used in canonical correlation	- SAT - GPA - Critical reading SAT score
Data analysis	Canonical correlation	Reader characteristic set	- Logical reasoning	- Logical Reasoning	- GALT
			- Factual chemistry knowledge	- Chemistry Content	- Multiple choice chemistry test
			- Working memory	- Working Memory	- Digit span test
			- Path length correlation	- Path Length Correlation	- Pathfinder
			- Nbrhd. similarity	- Neighborhood Similarity	- Pathfinder
		Reading comprehension set	- Encoding	- Text Recall	- Text recall
			- Lexical access	- Lexical Access	- Eye fixations
			- Bridging inferences	- Bridging Inferences	- MC bridging inference test
			- Elaborative inferences	- Elaborative Inferences	- MC elaborative inference test
			- General summary	- Delayed Recall	- Delayed recall of summary

Timeline

Session 1

The first meeting with the student was used to obtain relevant demographic information and three of the five reader characteristic variables (Logical Reasoning, Chemistry Content, and Working Memory). These questionnaires and tests were administered to the student in a group setting in a room where multiple students could complete them at the same time. Students were physically separated in the room to ensure that no personal information could be viewed by others in the group. The tests were administered via paper and pencil. The session lasted approximately one hour. When student consent was obtained, the overall SAT, SAT reading score, and high school GPA were obtained through the university.

Session 2

The second session was a one-on-one interview session held in the eye tracking lab in the chemistry department. This session was used to measure the two schema variables (Path Length Correlation and Neighborhood Similarity) and four of the five reading comprehension variables (Text Recall, Lexical Access, Bridging Inference, and Elaborative Inference). This session lasted approximately one hour. Each student was evaluated on the two chemistry concepts described above (Atoms, Ions, and Molecules, and Stoichiometry).

Schema measures.

When Session 2 started, students were instructed on how to complete a series of relatedness judgments on the two chemistry concepts presented. The instructions for this process are found in Appendix H. Before the student was asked to make relatedness judgments on concepts involved in the study, he/she was given a practice concept (Appendix H). This allowed the student to become familiar with the program. Once the student felt comfortable with the practice judgments he/she was given 105 word pair combinations per topic. This process resulted in a total of 210 relatedness judgments (105 per topic).

Reading comprehension measures.

After completing the Pathfinder program, each student was asked to read a series of chemistry passages (Appendix E). The passages covered the same two content areas as the Pathfinder program (Atoms, Ions, and Molecules, and Stoichiometry). Each topic had three passages resulting in a total of six chemistry passages. The chemistry passages consisted of one chemistry concept and included five to eight sentences on that concept. Each passage had a short title. Prior to reading each passage, the student was informed that once he/she had completed reading each passage, he/she would be asked to recall the passage. The student was also given a practice paragraph so that he/she could become familiar with the reading task. This task was not timed so the student was able to read the passages at the speed he/she naturally read for understanding. Once the student had recalled and typed as much of the previous passage as possible into the space provided on the computer, he/she progressed to the next passage until all six passages had been read and recalled. This procedure was completed using the eye tracking computer though the eye tracking data collected was only used during the student's reading of the passage, not during recall.

After reading each chemistry passage (three for each topic) and before progressing to the next paragraph, the student was asked to complete a multiple choice inference test. This test contained bridging and elaborative inference questions that included information necessary to understand each chemistry passage. These questions tested the student's ability to identify and make inferences during the reading of the chemistry passages. Session 2 was complete when the student had read the passage, recalled the text, and answered the inference questions for all six chemistry passages.

Session 3

A third session was held approximately 24-48 hours after session two and was used to measure the student's general summary of the passages. This session lasted approximately one hour and was completed using a computer. Eye tracking technology was not utilized during this session. In session three, the student was given the short title for each of the six chemistry passages to trigger his/her memory of the passage. The student was asked to provide summaries of the chemistry passage including what he/she felt were the main points in each passage. The scripts for this instruction are found in Appendix J. Participation in Session 3 was complete when the student had recalled a general summary of each passage.

Full participation in this study involved a 2-2.5 hour time commitment within seven days. The first and second session could occur within a one week period, but the third session had to occur approximately 24-48 hours after the second session. Table 8 provides a breakdown of the setting and tests that were administered to each student during the three sessions. This table only includes those variables that the student was tested on during the sessions, not the data for variables measured to ensure a wide range of students.

Table 8

Breakdown of Sessions

Session	**Setting**	**Variable**	**Test/Procedure used**	**Delivery Method**
Session 1	Group setting with individual testing	Chemistry Content	Multiple choice chemistry test	Paper and Pencil
		Chemistry background information and visual information	Demographic Questions	Paper and Pencil
		Logical Reasoning	Group Assessment of Logical Thinking	Paper and Pencil
		Working Memory	Digit Span Test	Electronic
Session 2: Should occur approximately within seven days of Session 1	Individual testing in eye tracking lab	Schema (Path Length Correlation and Neighborhood Similarity)	Pathfinder Program	Electronic (no eye tracking)
		Encoding	Recall typed on computer	Electronic (no eye tracking)
		Lexical Access	Fixation durations	Electronic (*eye tracking*)
		Inferences	Multiple choice inference test	Electronic (no eye tracking)
Session 3: Within 24-48 hours	Group setting with individual electronic testing	Situational Model	Recall of summary typed on computer	Electronic (no eye tracking)

Chapter 4 presents the results from the analysis of the data collected during these three sessions.

CHAPTER FOUR

This chapter presents the data and results of the study described in Chapter 3. The expert participant data and results will be presented first as these drove the selection of topics used in the student section of the study. The student participant data will be presented second including the testing of assumptions for each measure followed by the results obtained from the canonical correlations. Additional regression analysis will be presented as validation of the canonical correlation.

Expert Participants

The expert portion of the study was conducted for two purposes. The first was to determine which two of the four topics initially selected by the researcher would produce the best Pathfinder networks. This selection was based on three measures: coherency; path length correlations; and neighborhood similarities. The two topics that performed the best on these three measures were then used in the remainder of the study with the student participants. Only the top two topics could be used due to the time such a methodology requires. The analysis of a third topic would require too high a demand on students' time and cognitive effort. The second purpose was to develop a referent expert Pathfinder network for each of the two chemistry topics by averaging all expert networks together. The two averaged networks (Atoms, Ions, and Molecules and Stoichiometry) were used as referent networks for comparison with the student-derived Pathfinder networks to determine the quality of each student's network.

The selection of two topics for use in the student portion of the study was determined using the results from the experts' Pathfinder data in the following methodology. Four chemistry topics were selected by the researcher for the expert participant portion of this study including the following: Atoms, Ions, and Molecules; Stoichiometry; Chemical Bonding; and Thermodynamics. These topics were chosen to represent a range of chemistry topics taught in general chemistry courses. Sixteen key terms were created for each topic. These were selected by cross referencing the list of vocabulary words found in several widely used general chemistry texts in the corresponding chapter for each of the four topics (Brown, LeMay, Bursten, & Burdge, 2003; Kotz, Treichel, & Weaver, 2006; Silberberg, 2000). The resulting lists contained between 20 and 30 key terms which were then reduced to the target number of 15. Fifteen key terms results in 105 relatedness judgments made by each expert. This number has been shown in previous studies to be the lowest number of key terms and relatedness judgments that results in accurate Pathfinder networks (Clariana, 2003). This reduction was accomplished by removing some supplementary sections of the topic's chapter that were not critical to the main topics presented in the chapter. The set of topics and lists of key terms were reviewed by two chemistry instructors who taught general chemistry at the high school and undergraduate level. These instructors verified that the key terms used were both important to each topic and represented a wide range of important concepts in the topic. The lists used for each topic are in Appendix D.

Seven chemistry experts were included in this study. These experts ranged from undergraduate chemistry professors at a midsize private institution to chemists currently working in chemical research. The experts were asked to make relatedness judgments in the Pathfinder program for each list of key terms as described in Chapter 3. For three of the topics, all seven experts completed the ratings. In the final topic, thermochemistry, one expert was not able to complete the ratings due to time constraints. The ratings were then used to create a Pathfinder network (PFNET) per topic for each expert. The following data analysis of the experts' networks were used in the selection of the two best performing topics from the original four.

Selection of Two Best Performing Topics

Coherency.

To determine which topics were best for use with the student population, each expert's network was first checked for coherency using the Pathfinder program. Coherency is a reflection of the consistency of the data. The coherency of a set of proximity data is based on the assumption that the relatedness between a pair of items can be predicted by the relationships of the items to other items in the set. Very low coherency (below 0.2) may indicate that the participant is not an expert in that content area and should not be used as such (Pathfinder 6.3, 2011). A coherency score below this level was deemed unacceptable. The coherency ranges for each topic are presented in Table 9.

Table 9

Expert Coherency Scores[a]

Topic	Min. Coherency	Max. Coherency	Experts with Low Coherency
Atoms, Ions, and Molecules N = 7	0.50	0.71	0
Stoichiometry N = 7	0.25	0.75	0
Chemical Bonding N = 7	0.19	0.67	1
Thermodynamics N = 6	0.04	0.58	2

[a] Coherency scores can range from zero to one.

Two experts' ratings were found to be below the necessary level for coherency on the Thermodynamic topic. One of these experts was also below the coherency threshold (0.20) on the Chemical Bonding topic. These experts' PFNETs were removed from the topics in which their performance was low. The Thermodynamic topic had two experts with low coherency values. This coupled with the fact that only 6 experts had completed this topic resulted in only four acceptable expert networks. This topic was therefore not chosen as one of the two topics to be further used in this study. In the Chemical bonding topic the expert's PFNET with low coherency was removed for the remainder of the analysis resulting in the chemical bonding topic having an n = 6.

Averaging expert networks.

The expert PFNETs were averaged by the Pathfinder program to produce an overall average expert network for each topic. The program can create visual representations of the mathematical relationships found by the program. Figures 4-6 show the visual representations of the average expert network for each topic.

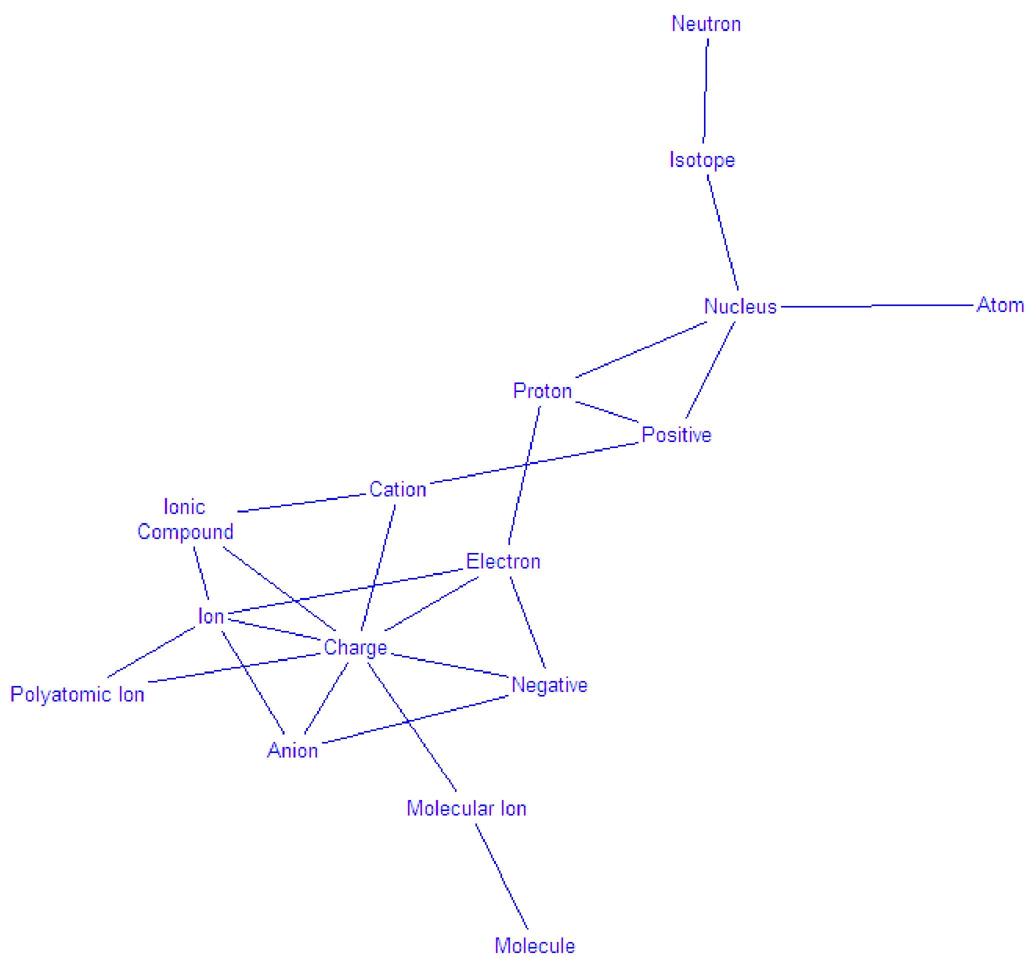


Figure 4. Visual representation of the average expert Atoms, Ions, and Molecules network derived by the Pathfinder program.

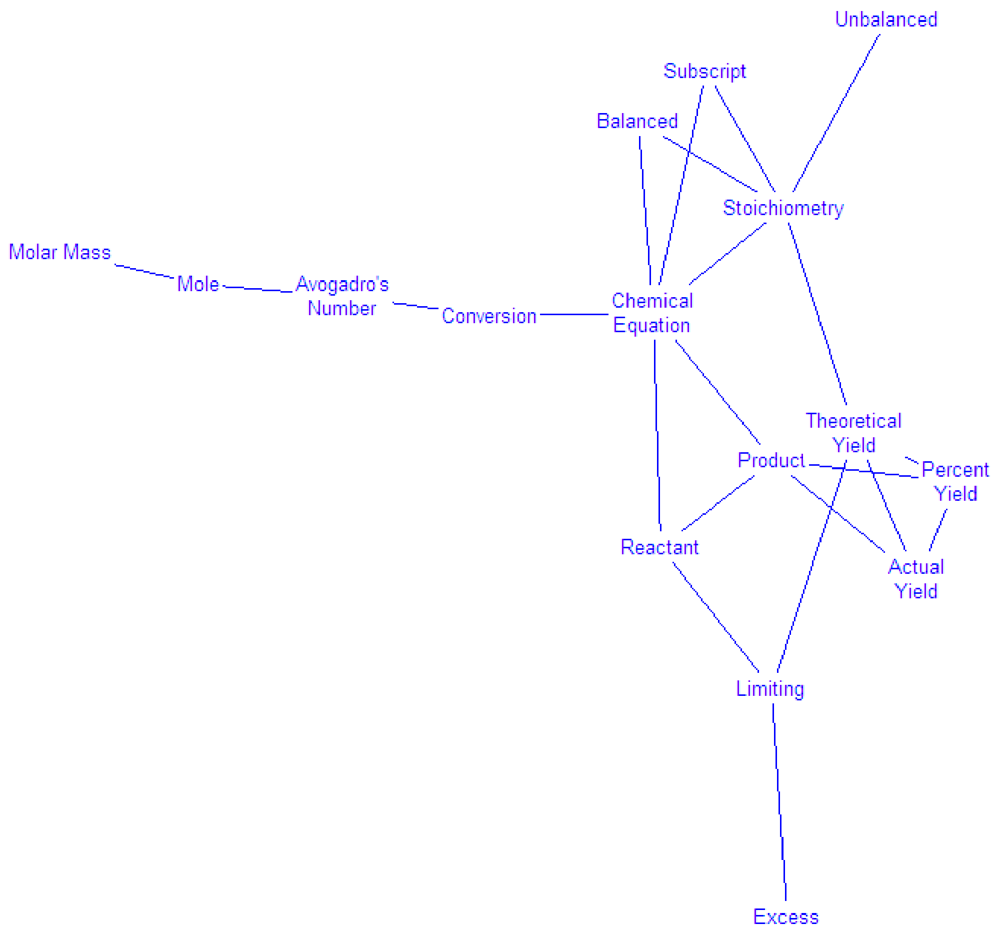


Figure 5. Visual representation of the average expert Stoichiometry network derived by the Pathfinder program.

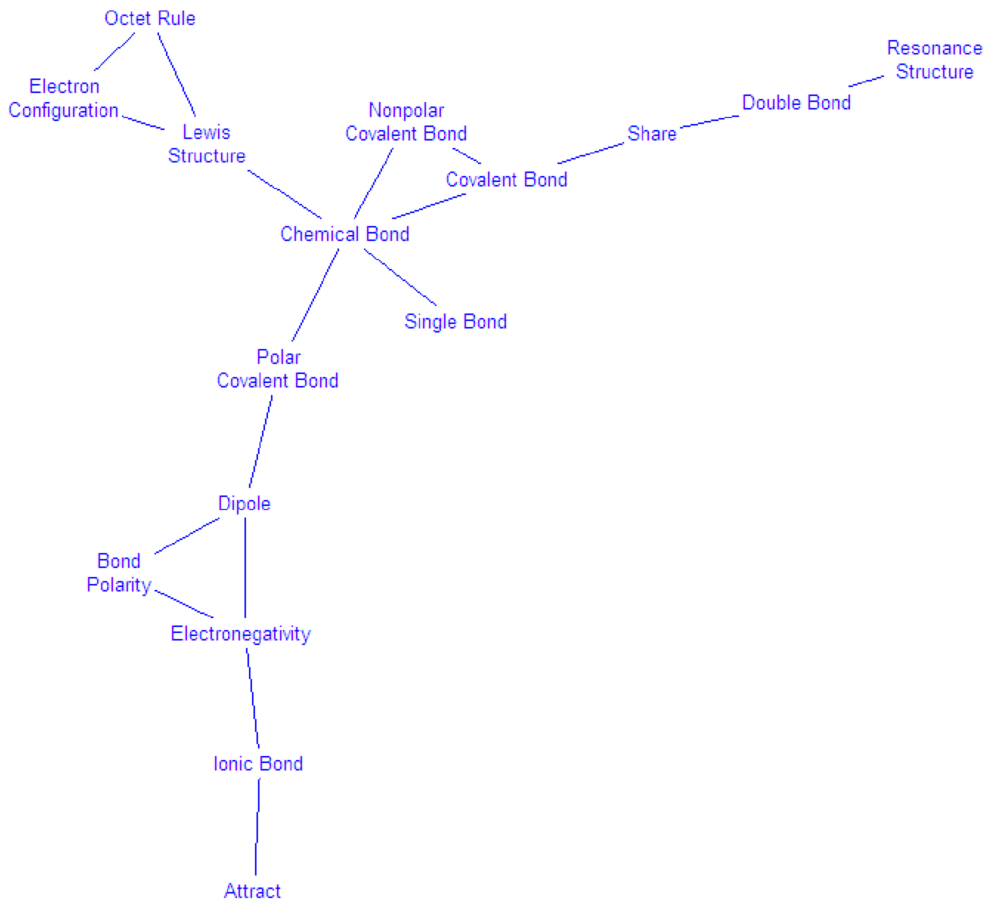


Figure 6. Visual representation of the average expert Chemical Bonding network derived by the Pathfinder program.

These figures show a visual representation of the relationships found by the Pathfinder program of how the experts, on average, organized the key terms in their minds for each chemistry concept. The Pathfinder program then mathematically compares the path lengths and neighborhood similarities of these networks to determine how similar two networks are.

Each individual expert network was analyzed to determine the statistical similarity between the individual network and the average expert network for each topic. This was done to ensure that no one expert would unduly influence the averaged network's structure.

The networks were compared on two measures path length correlations and neighborhood similarities. These are the same two measures that were later used to compare each student's PFNET to the average expert PFNET.

Path length correlations.

Each expert's network was compared to average experts' network on two measures, path length correlations and network similarity. The first measure, path length correlation, is similarity of path lengths between the expert's individual network and the referent experts' network. This is determined by measuring the ratio of shared attributes of the two networks by determining the correlation coefficient of the networks' path lengths. The method for determining the coefficient was discussed previously in Chapter 3. Table 10 presents the correlation between path lengths of each expert and the average expert network. This is presented for each topic as well as the average correlation per each topic.

Table 10

Expert Path Length Correlations[a]

Expert ID	**Atoms, Ions, and Molecules**	**Stoichiometry**	**Chemical Bonding**
001	0.833	0.694	0.616
002	0.794	0.727	0.704
003	0.722	0.638	0.592
004	0.893	0.690	0.609
005	0.736	0.627	0.651
006	0.879	0.618	0.761
007	0.744	0.640	x
Average Correlation	0.800	0.662	0.656

[a]Path length correlations can range from zero to one.

This table shows that all expert/average expert correlations are large based on the correlation values generally accepted as small (r = 0.10 to 0.29), medium (r = 0.30 to 0.49), and large (r = 0.50 to 1.0) (Cohen, 1988). This means that based on this analysis, no one expert was unduly affecting the overall average network. The two topics with the highest average correlations are Atoms, Ions, and Molecules with r = 0.800, and Stoichiometry with r = 0.662.

Neighborhood similarities.

The second measure used to assess the similarity of the networks was a measure of the neighborhoods, or clusters of nodes, found within the two networks' corresponding nodes. As stated in Chapter 3, this measure is referred to as the neighborhood similarity. The neighborhood similarity found between each expert's network and the average expert network, along with the averages per topic is shown in Table 11.

Table 11

Expert Neighborhood Similarities[a]

Expert ID	**Atoms, Ions, and Molecules**	**Stoichiometry**	**Chemical Bonding**
001	0.441	0.556	0.364
002	0.270	0.600	0.293
003	0.457	0.292	0.276
004	0.485	0.567	0.357
005	0.419	0.306	0.256
006	0.528	0.421	0.324
007	0.448	0.324	x
Average Correlation	0.435	0.438	0.312

[a] Neighborhood Similarities can range from 0 to 1.

This table shows that the two topics with the highest average neighborhood similarities are the Atoms, Ions, and Molecules topic with an average neighborhood similarity value = 0.435, and the Stoichiometry topic with an average neighborhood similarity value = 0.438. These topics show neighborhood similarities comparable to those reported as high in the literature in Pathfinder network neighborhoods (Schvaneveldt et al. 1985).

The three topics were compared on the three measures discussed (coherency, path length correlation, and neighborhood similarity) to determine which would be most appropriate for use in the student section of this study. Table 12 presents this comparison.

Table 12

Comparison of Topics to Determine Top Performing Two

Topic	**Number of participants with appropriate coherency**	**Average Path Length Correlation Coefficients**	**Average Neighborhood Similarities**
Atoms, Ions, and Molecules	7	0.800	0.435
Stoichiometry	7	0.662	0.438
Chemical Bonding	6	0.656	0.312

The topics Atoms, Ions, and Molecules and Stoichiometry performed best in the number of experts with high coherency values, their average correlation coefficients, and their average neighborhood similarity values. Based on the results, these two topics were chosen for use in the student portion of the study.

Creating referent expert networks.

The average experts' network of the two topics selected (Atoms, Ions, and Molecules, and Stoichiometry) were also used as the referent networks for the student section of this study. These average referent experts' networks were used to determine the quality of each student's Pathfinder network. As stated above, the expert section of this study was used to determine the best two topics and create an averaged referent expert network for each topic to be used in the student portion of this study. Once the two topics were selected and the referent experts' network were created, the researcher could proceed with the student portion of this study.

Student Participants

Two hundred and seventy four students enrolled in chemistry classes were invited to participate in this study and 85 agreed. Of the 85 who agreed, forty-three students attended the sessions and participated in this study. Table 13 presents the chemistry courses the students were currently enrolled in.

Table 13

Breakdown of Student Participants from Each Chemistry Course

Chemistry Course	Number of Students
Non-science Major's Chemistry	7
Nursing Chemistry	7
Engineering and Major's General Chemistry	12
Organic Chemistry	7
Upper Level Chemistry (Physical Chemistry and Analytical Chemistry)	10

The overall SAT scores, reading comprehension sub-scores, and high school GPA were obtained, with student permission, from the university. The high school GPA and SAT overall score were weighted at 50% each to give equal value to both scores and were then combined into one predictive college success score. As discussed in Chapter 3, the combination of these scores provides the best prediction of student success in their undergraduate degree. The combined SAT and GPA scores were plotted to determine if they represented a wide range of ability. This was also done with the SAT reading subscore. A normal distribution was found on both scores (combined GPA/SAT and SAT reading subscore). These results indicate that the sample of students is normally distributed on these variables.

Reader Characteristic Set

The Forward Digit Span Test, Group Assessment of Logical Thinking (GALT) and chemistry content test were used to determine the students' score on the working memory, logical reasoning, and factual chemistry knowledge variables. As reported in Chapter 2, these variables are expected to have an effect on different aspects of reading comprehension when the text is scientific in nature. Table 14 shows the range, mean, and standard deviation of the scores of each variable.

Table 14

Descriptive Statistics of Working Memory, Logical Reasoning, and Chemistry Content

Variable	Range of Responses	Mean	Standard Dev.
Forward Digit Span Test[a]	5-9	7.16	0.90
GALT[b]	3-12	9.23	2.29
Chemistry Content Test[c]	4-19	12.34	5.0

[a]Forward Digit Span Test scores can range from 0 to the upper limit demonstrated by the participant.
[b]GALT scores can range from 0 to 12.
[c]Chemistry Content Test scores can range from 0 to 20.

Each of these variables displayed a normal distribution of scores.

Using the Pathfinder program, the PFNETs were analyzed to determine the similarity between the students' PFNET and the referent experts' network for both topics, Atoms, Ions, and Molecules, and Stoichiometry. Similarity was again determined using Path Length Correlation and Neighborhood Similarities. These values could range from 0—1 with 1 representing a student network that was very similar to the referent experts' network. The scores were averaged for each student across the two chemistry topics. Table 15 displays the range of values, means, and standard deviation of the scores for these two variables.

Table 15

Students' Descriptive Statistics of Two Measures of PFNet Similarity

Variable	Mean	Standard Dev.	Range
Path length correlation	0.503	0.216	0.109—0.815
Neighborhood similarity	0.326	0.105	0.152—0.556

These PFNet similarity variables displayed a normal distribution of scores.

Reading Comprehension Set

The Text Recall, Bridging Inference questions, and Elaborative Inference questions were graded resulting in a score for each variable per chemistry passage. As reported in Chapter 2, these important variables for reading comprehension when the text is scientific in nature. These scores were averaged across the two chemistry topics. Table 16 displays the range of values, means, and standard deviation of the scores for these three variables.

Table 16

Descriptive Statistics of Text Recall, Bridging Inference, and Elaborative Inference Variables

Variable	**Mean**	**Standard Dev.**	**Range**
Text Recall	41.52	16.19	1.96—72.14
Bridging Inferences	7.17	1.86	2.00—9.00
Elaborative Inferences	7.00	1.76	2.00—9.00

The scores on each of these variables displayed a normal distribution.

As discussed in Chapter 3, the students' lexical access ability was measured using eye tracking technology. The eye tracking technology records students' fixation durations. Specifically, the students' fixation durations on scientific words were measured as an indicator of the lexical access ability. These durations were then divided by the number of letters in each scientific word resulting in fixation duration per letter score for each scientific word. As shown in the studies presented in Chapter 2, larger scores indicate that the students needed a longer amount of time to access the relevant lexical information on the scientific words. Each student's Lexical Access scores were averaged across all scientific words in both topics resulting in one average lexical access score. This score was then subtracted from one to result in larger scores reflecting better lexical access (faster times) and lower scores reflecting poor lexical access (slower times). This was done so that all variables had a positive correlation for the analysis. Table 17 presents the range, mean, and standard deviation of the Lexical Access scores.

Table 17

Descriptive Statistics of Lexical Access Variable

Variable	**Mean**	**Standard Dev.**	**Range**
Lexical Access	0.827	0.105	0.337—0.928

The scores for lexical access were found to have a normal distribution.

The students' ability to remember a general summary of the chemistry passages was measured using a Delayed Recall task. The Delayed Recall task for both chemistry topics were graded by the researcher using a validated rubric as described in Chapter 3. The set of students' responses along with the relevant rubric for each topic were sent to two outside graders—one for each topic. The scores determined by the researcher and outside reviewer were analyzed to determine inter-rater reliability in the grading. Table 18 presents the results of this analysis.

Table 18

Inter-Rater Reliability of Delayed Recall Scoring

Raters	t	sig.
Rater 1 and Rater 2	0.485	0.629
Rater 1 and Rater 3	0.487	0.628

These results indicated that there was no significant difference in the grading of the researcher and either one of the two outside graders. The researcher's score and the outside grader's score for each student participant were averaged between the graders, across the chemistry passages, and across the chemistry topics. The scores of the graders were averaged to include multiple perspectives in the grading. This resulted in one Delayed Recall grade for each student. Table 19 presents the range, mean, and standard deviation of the Delayed Recall scores.

Table 19

Descriptive Statistics for the Delayed Recall Variable

Variable	Range	Mean	Standard Dev.
Delayed Recall	0.44—13.00	6.251	3.147

The scores in the Delayed Recall variable displayed a normal distribution.

Assumptions Testing

Prior to running the canonical correlation for the main analysis of this study, each variable and set of variables was investigated for violations of the assumptions relevant to canonical correlations. These assumptions include sample size, outliers, homoscedasticity, normality, linearity, and independence of variables.

Sample Size and Outliers

If a multivariate regression is performed on a sample that is too small, the results may not be generalizable to other samples. Recommendations for sample size vary widely depending on the reference. A common recommendation is 10-20 participants per predictive variable (Stevens, 1996, Tabachnick and Fidell, 1996). D'Agostino and Stephens (1986) states that strong regressions ($R > 0.7$) can be detected using much smaller n's. The strength of the canonical correlation found in this study, which will be presented below, is very strong. This means that an n of 41 was adequate for the level of strength of the canonical correlation found in this study. The data were tested for outliers, which are those data points with standardized residual values above $|3.3|$ (Tabachnick and Fidell, 1996). This was done using the SPSS program's Mahalanobis distances. Few outliers were found and all were deemed to be the result of input error. These mistakes were rectified and the data were determined to be appropriate for statistical analysis based on both the sample size and outliers assumptions.

Homoscedasticity, Normality, Linearity, and Independence of Residuals

Residuals are the differences between the obtained and the predicted dependent variable score. Residuals are the values that are tested for assumptions in a canonical correlation analysis. There is no process by which the homoscedasticty, normality, linearity, or independence of residuals in a multivariate set can be determined as a whole, so it is common practice to investigate each individual variable and use that as an indicator of the set's appropriateness for the statistical analysis. Each variable's residual distribution was plotted and analyzed. The first assumption, Homoscedasticity, involves the variance of residuals around the predicted dependent variable scores. These values should be approximately the same for all the predicted scores. The residual distributions were evaluated and found to be appropriate for the variables used in this study. The variables in this study also displayed a normal residual distribution and thus were found not to violate the normality assumption. Each variable was also plotted one at a time against every other variable to check for linearity. No two variables displayed a non-linear relationship and thus all variables were determined appropriate for analysis on the linearity assumption.

Independence of residuals ensures that no two variables are dependent on one another. This is done through research of the literature and development of the research methodology. Based on the literature reviewed in Chapter 2, the variables involved in this study were selected so that no two were dependent on one another. For this reason, the assumption of independence of variables is a valid assumption for this study.

Canonical Correlation Analysis

The variables involved in the canonical correlation consisted of two sets, the Reader Characteristic set that is considered the independent set and the Reading Comprehension set which is the dependent set. Table 20 lists all of the variables involved in the analysis.

Table 20

Variables Involved in the Canonical Correlation

Reader Characteristic Set (Independent)	**Reading Comprehension Set (Dependent)**
Working Memory	Text Recall
Logical Reasoning	Lexical Access
Chemistry Content	Bridging Inference
Path Length Correlation	Elaborative Inference
Neighborhood Similarity	Delayed Recall

To better understand the canonical correlation analysis it may be helpful to review the following terms (Hair, 1998).

Canonical Correlation: Measure of the strength of the overall relationships between the linear composites (*canonical variates*) for the independent and dependent variables.

Canonical Variates: Linear combinations that represent the weighted sum of two or more variables. Canonical variates can be defined for either dependent or independent variables.

Canonical Function: Relationship (correlational) between two linear composites (*canonical variates*). Each canonical function has two canonical variates, one for the set of dependent variables and one for the set of independent variables.

Canonical Roots: Squared canonical correlations, which provide an estimate of the amount of shared variance between the respective optimally weighted canonical variates of dependent and independent variables.

To determine the correlations between the various variables, a correlation matrix was created for each set of variables. These matrices are presented in Table 21 and 22.

Table 21

Correlation Matrix Reader Characteristic Set

	Working Memory	**GALT**	**Chemistry Content**	**Path Length Correlation**	**Neighborhood Similarity**
Working Memory	1.000	0.085	0.019	0.220	0.216
GALT	0.085	1.000	0.497	0.559	0.541
Chemistry Content	0.019	0.497	1.000	0.647	0.699
Path Length Correlation	0.220	0.559	0.647	1.000	0.840
Neighborhood Similarity	0.216	0.541	0.699	0.840	1.000

Table 22

Correlation Matrix Reading Comprehension Set

	Text Recall	Lexical Access	Bridge Inference	Elaborative Inference	Delayed Recall
Text Recall	1.000	0.302	0.548	0.475	0.614
Lexical Access	0.302	1.000	0.072	0.034	0.250
Bridge Inference	0.548	0.072	1.000	0.747	0.562
Elaborative Inference	0.475	0.034	0.747	1.000	0.556
Delayed Recall	0.614	0.250	0.562	0.556	1.000

Evaluation of the Canonical Correlation

The canonical correlation analysis was restricted to deriving five canonical functions which is based on the lowest number of variables in either set. These functions relate the linear regression of one set of variables to the linear regression of the other. In this study only the first function was evaluated as it was the only function that significantly accounted for the variance in the variables. This is the function that maximizes the variance between the two sets of variables. The function is presented in Table 23.

Table 23

Canonical Correlation

Root No.	Eigenvalue	Percent	Canonical Correlation (R)	Squared Correlation (R^2)
1	3.597	84.78	0.8856	0.7825

The significance of the values in this table will be evaluated in detail next through an investigation of the following: (1) the level of statistical significance of the function, (2) the magnitude of the canonical correlation, (3) the redundancy index for the percentage of variance accounted for from the two data sets. Taking all three of these analyses into account when evaluating the model will provide an in-depth description of the effectiveness of this model in describing the variance present in student responses.

Statistical significance of the function.

The significance of the canonical correlation is an evaluation of each canonical root determined by the SPSS program. The canonical root provides an estimate of the shared variance between the optimally weighted dependent and independent variables. This estimate is tested for significance using discriminant functions including Pillai's trace, Hotelling's trace, Wilk's trace, and Roy's gcr. The results of this analysis are presented in Table 24.

Table 24

Multivariate Tests of Significance

Discriminant Function	**Value**	**df**	**F**	**Sig.**
Pillai's trace	1.252	(25, 175)	2.34	0.001
Hotelling's trace	4.240	(25, 147)	4.99	0.000
Wilk's	0.126	(25, 117)	3.48	0.000
			Critical Value	**Sig.**
Roy's gcr	0.782	(s = 5, m = -1/2, n = 14 ½)	$p_{(0.05)}$= 0.399 $p_{(0.01)}$= 0.472	Significant at 0.01 level

The values in this table show that all multivariate tests of significance find the function to be significant. This means that the canonical function accounts for a statistically significant portion of the variance in student responses on the variables involved in this study.

Magnitude of the canonical correlation.

The second consideration of the effectiveness of the canonical function to describe the variance in student responses is the magnitude of the canonical correlation. This is a practical measure of the size of canonical correlation. No standard measure or guidelines have been established to determine the necessary size of a meaningful canonical correlation. The decision of whether or not the canonical correlation is large enough is generally based on the contribution of the findings to the research field of the study. The magnitude of the canonical correlation in this study was determined by evaluating the squared correlation (R^2), also known as the canonical root because it provides an estimate or percentage of the amount of shared variance between the optimally weighted dependent and independent variables. In this study the canonical root of the significant function was 0.782. This means that 78.2% of the variance found in student responses is accounted for by the canonical correlation function. When compared to other studies in the field of education and scientific literacy, which typically range from 0.300 to 0.600, the percentage reported here is large (Ozkal, 2011, Turmo, 2004).

Redundancy measure of shared variance.

The squared canonical correlations (roots) provide an estimate of the shared variance between the canonical variates. This measure of shared variance can sometimes lead to misinterpretation. The misinterpretation can occur when the squared canonical correlations represents the variance shared by the linear composites of the sets of dependent and independent variables, and not the variance extracted from the sets of variables (Hair et al, 1998). Thus, a relatively strong canonical correlation may be obtained between the two sets of variables even though the linear composites may not extract significant portions of variance from their respective sets of variables. To avoid this misinterpretation the redundancy index of the canonical correlation was evaluated.

The redundancy index is a summary measure of the ability of a set of variables to explain variation in each individual variable in the opposite set. It involves computing the squared multiple correlation coefficients between the total independent variable set and each variable in the dependent variable set. These squared correlations are then averaged and multiplied by the squared canonical correlation. This process is outlined in Table 25.

Table 25

Calculation of Redundancy Index

Variable	**Correlation**	**Correlation Squared**	**Average Correlation Squared**	**Canonical Correlation (R^2)**	**Redundancy Index**
Independent Set					
Working Memory	0.070	0.005			
GALT	0.606	0.368			
Chemistry Content	0.838	0.701	0.550	0.7825	0.430
Path Length Correlation	0.986	0.973			
Neighborhood Similarity	0.837	0.701			
Dependent Set					
Text Recall	0.652	0.426			
Lexical Access	0.236	0.056			
Bridging Inferences	0.822	0.676	0.524	0.7825	0.410
Elaborative Inferences	0.926	0.856			
Delayed Recall	0.779	0.607			

When compared to other studies in the field of chemical education and scientific literacy that use canonical correlation analysis, the values found here are within the acceptable range of redundancy index values (Ozkal, 2011, Turmo, 2004). Compared to these studies the redundancy of the independent set (0.430) and the dependent set (0.410) is in the mid range. There is no set minimum value accepted as the lowest acceptable redundancy measure (Hair et al, 1998). It should be noted both working memory and lexical access have low redundancy with the opposite set of variables. This means that they have a low ability to explain the set of variables of which they are not a part. This phenomenon will be evaluated in more detail in the next section of this chapter.

The canonical correlation function in this study evaluated the three criteria described above (statistical significance, magnitude, and redundancy). The canonical function performed well on all three criteria. With the canonical relationship deemed statistically significant and the magnitude of the canonical root and the redundancy index acceptable, further interpretations of the results will follow.

Evaluation of the Variables

Further interpretation of the canonical correlation analysis involves examination of the canonical function to determine the relative importance of the original variables in the two sets of variables in canonical correlation. This examination will include the following criteria: (1) statistical significance of the separate regression of each independent variable, (2) canonical weights (standardized coefficients); (3) redundancy; and (4) structure correlations.

Separate regressions of each independent variables.

The separate regressions of each independent variable were evaluated to determine which variables were contributing significantly to the explanation of the variance in the dependent set of variables. Table 26 presents the results of these regressions.

Table 26

Separate Regressions of the Independent Variables

Variable	R^2	**F**	**Sig.**
Working Memory	0.091	0.700	0.627
Logical Reasoning	0.311	3.156	0.019
Chemistry Content	0.568	9.204	0.000
Path Length Correlations	0.764	22.674	0.000
Neighborhood Similarities	0.632	12.038	0.000

This table shows that all reader characteristic variables except working memory describe a significant amount of the variance in the Reading Comprehension set of variables. The fact that the Working Memory variable does not describe a significant amount of variance in the Reading Comprehension set will be discussed in the limitations section of Chapter 5. The R^2 provides the percentage of variance each variable describes. Thus Logical Reasoning describes 31.1% of the variance, Chemistry Content describes 56.8% of the variance, Path Length Correlations describes 76.4% of the variance and Neighborhood Similarities describes 63.2% of the variance. The variance described obviously overlaps on some of these variables but the order of importance, as determined by the R^2, is the following:

1) Path Length Correlations
2) Neighborhood Similarities
3) Chemistry Content
4) Logical Reasoning

This shows that the two Pathfinder measures of schema describe the most variance. This result will be validated by later regression analysis.

Canonical weights.

Another approach to interpreting canonical functions involves examining the magnitude of the canonical weight associated with each variable. Variables with larger weights contribute more to the variates then those with lower weights. To interpret the relative importance of the canonical weights, they must first be standardized so that the comparison can be made. Table 27 contains the standardized canonical weights for the dependent (reading comprehension) and independent (reader characteristic) variables used in this study.

Table 27

Standardized Canonical Weights of Independent and Dependent Variables

Variable	Standardized Canonical Coefficient
Independent Set	
Working Memory	0.158
Logical Reasoning	0.058
Chemistry Content	0.033
Path Length Correlation	0.955
Neighborhood Similarity	0.074
Dependent Set	
Text Recall	0.003
Lexical Access	0.177
Bridging Inference	0.163
Elaborative Inference	0.655
Delayed Recall	0.281

As stated above, the magnitude of the weights represents the variable's relative contribution to the variate. Based on the size of the weights, the order of contribution of variables to each set is as follows:

Reader Characteristic Set (Dependent):

1. Path Length Correlation
2. Working Memory
3. Neighborhood Similarity
4. Logical Reasoning
5. Chemistry Content

Reading Comprehension Set (Independent):

6. Elaborative Inferences
7. Delayed Recall
8. Lexical Access
9. Bridging Inference
10. Text Recall

The order of these variables indicates which variables have the greatest effect on the canonical correlation. It can be seen here that the variable Path Length Correlation has the greatest effect on the canonical correlation (0.995). Elaborative Inference (0.655) and Delayed Recall (0.281) explain the greatest amount of variance of the dependent variables. The other seven variables account for very small amounts of the variance (0.17 or less). These results indicate that the Path Length Correlation, Elaborative Inference, and Delayed

Recall variables are contributing the greatest amount to the variance described by the canonical function.

Redundancy.

The redundancy of a variable is its ability to describe the set of variables of which it is not a part. This is different from the redundancy index described previously as it is associated with each individual variable not the set as a whole. The redundancy is important in determining each variable's relative importance to the canonical function. No set range of redundancy values has been set as low, medium, or high. One study found in the literature report values around 0.200 low (Huba and Bentler, 1980). Table 28 provides the redundancy values for each variable.

Table 28

Redundancy Values of Variables[a]

Variable	Redundancy
Independent Set	
Working Memory	0.004
Logical Reasoning	0.288
Chemistry Content	0.549
Path Length Correlation	0.761
Neighborhood Similarity	0.549
Dependent Set	
Text Recall	0.333
Lexical Access	0.044
Bridging Inferences	0.529
Elaborative Inferences	0.700
Delayed Recall	0.475

[a]Redundancy values can range from zero to one.

Medium to high redundancy is desired as that is interpreted as the variable describing a large percent of the variance in the opposite set of variables. In the reader characteristic (independent) set of variables, Chemistry Content, Path Length Correlation, and Neighborhood Similarity have medium to high ranges and therefore each describe a good amount of the variance in the reading comprehension variables. Working Memory and Logical Reasoning have low redundancy values and thus do not describe a large amount of the variance.

In the reading comprehension (dependent) set of variables only the two inference variables and Delayed Recall have medium to large redundancy scores. This means that these three variables describe a medium to large amount of the variance found in the Reader Characteristic set of variables. Text Recall and Lexical Access would be described as demonstrating small redundancy scores and thus do not describe a large amount of the variance of the Reader Characteristic set. In summary, Chemistry Content, Path Length Correlation, Neighborhood Similarity, Bridging and Elaborative Inferences, and Delayed Recall each have a high redundancy index and play an important role in the description of variance in the canonical correlation.

Structure correlations.

Structure correlation measures the simple linear correlation between each variable in the dependent or independent set and the set's canonical variate. This reflects the variance that the observed variable shares with the canonical variate and can be used in assessing the relative contribution of each variable to the canonical function. Table 29 presents the structure correlations for each variable.

Table 29

Structure Correlations Between Each Variable and the Set's Contribution to the Canonical Variate

Variable	Structure Correlation
Independent Set	
Working Memory	0.071
Logical Reasoning	0.606
Chemistry Content	0.838
Path Length Correlation	0.986
Neighborhood Similarity	0.837
Dependent Set	
Text Recall	0.652
Lexical Access	0.236
Bridging Inference	0.822
Elaborative Inference	0.926
Delayed Recall	0.779

Four of the five independent variables, Logical Reasoning, Chemistry Content, Path Length Correlation, and Neighborhood Similarity, have high structure correlation according to the generally accepted values of small ($r = 0.10$ to 0.29), medium ($r = 0.30$ to 0.49), and large ($r = 0.50$ to 1.0) (Cohen, 1988). This indicates that these variables have a high correlation between their scores and the canonical variate of the set they are a part of (Reading Comprehension Set). When analyzing the dependent set of variables, Logical Reasoning, Chemistry Content, Path Length Correlation, and Neighborhood Similarity have high correlations between their scores and the canonical variate of the set they are a part of (Reader Characteristic set). Working Memory and Lexical Access have little to no correlation with their set's canonical variate. This could mean that Working Memory does not greatly affect the scores on any of the dependent variables (Reading Comprehension) and that Lexical Access is not affected by any of the independent variables (Reader Characteristic). These results cause some limitations to the interpretation of the canonical correlation analysis, which will be further discussed in the limitations section of Chapter 5.

Overview of Canonical Correlation

The results from this analysis indicate that there is indeed a relationship among the variables involved in this study and that the relationship is significant. This was determined by the fact that the canonical function was significant, the magnitude of the function was large (78.2% of the variance described), and the redundancy index of each set was acceptable when compared to other studies in relevant fields (Independent= 0.430 and Dependent = 0.410). For these reasons, it was determined that the relationships between the variables involved in this study were indeed significant and therefore appropriate for further investigation.

The individual variables were evaluated using four analyses: (1) the statistical significance of the separate regressions of the independent (Reader Characteristic) variables; (2) the canonical weights of the variables; (3) the redundancy of the variables; (4) the correlation between each variable and the canonical function (structure correlation). These four analyses will now be compared to determine which variables play the most important roles in the canonical correlation function.

The separate regressions for each of the independent variables, Working Memory, Logical Reasoning, Chemistry Content, Path Length Correlation, and Neighborhood Similarity, were compared to determine which were significant, and of those that were significant, which accounted for the most variance in the dependent (Reading Comprehension) set of variables. The Working Memory regression was not significant and therefore this regression did not account for a significant amount of the variance in the Reading Comprehension set of variables. The remaining variables' regressions were significant. The percent variance accounted for by each variable was the following in order of greatest to least:

1. Path Length Correlation (76.7%)
2. Neighborhood Similarity (63.2%)
3. Chemistry Content (56.8%)
4. Logical Reasoning (31.1%)

This shows that the two Pathfinder measures provide the highest percentage of variance explained, followed by Chemistry Content and Logical Reasoning variables.

The canonical weights were compared next to determine which variables were weighted highest in the canonical correlation and thus most important to the function. The weights for both sets of variables are listed from greatest to least:

Reader Characteristic Set:

1. Path Length Correlation
2. Working Memory
3. Neighborhood Similarity
4. Logical Reasoning
5. Chemistry Content

Reading Comprehension Set:

1. Elaborative Inferences
2. Delayed Recall
3. Lexical Access
4. Bridging Inferences
5. Text Recall

It is surprising in the Reader Characteristic set that Working Memory rates high on the list of weights when the separate regressions have shown that this variable does not have a significant regression. It is possible the weight of this variable comes from its low correlation with the canonical function.

Redundancy of each variable was evaluated to determine the amount of variance each variable described in the opposite set of variables. In the Reader Characteristic set of variables, Path Length Correlation, Neighborhood Similarity, and Chemistry Content were found to have medium to large redundancies (from largest to smallest). In the Reading Comprehension set of variables, Elaborative Inferences, Bridging Inferences, and Delayed Recall were found to have medium to large redundancies (from largest to smallest). These variables described the greatest amount of variance in the opposite set.

Finally, structure correlations were evaluated to determine which variables were highly correlated, and therefore predictive, of the canonical function. In each set the following variables were found to have medium to large correlations ($r = 0.30$-1.00) in order from largest to smallest:

Reader Characteristic Set

1. Path Length Correlation
2. Chemistry Content
3. Neighborhood Similarity
4. Logical Reasoning

Reading Comprehension Set

1. Elaborative Inference
2. Bridging Inference
3. Delayed Recall
4. Text Recall

In both sets, one variable was found to have a small correlation (Working Memory and Lexical Access). This is interpreted as these variables not being strongly correlated with the canonical function.

The analyses for individual variables will be compared at the end of this chapter to determine which variables have the greatest effect on the canonical correlation. Prior to that evaluation, a validation of the canonical correlation findings is presented.

Validation of Canonical Correlation Findings

Separate regression analyses were performed to validate the findings of the canonical correlation as well as to further investigate the relationships between the independent and dependent variables. A separate regression was run with the five reader characteristic variables as predictor variables and each reading comprehension variable as the dependent or outcome variable. Since all ten variables met the assumptions for a canonical correlation analysis, they also are appropriate for regression analysis. The results of these separate regressions will now be discussed.

The regressions were run with all five independent variables in the model to see which variables are significantly contributing to the variance in each dependent variable. Table 30 provides the percent of variance described by the regression model (R^2), the F value, and significance for each dependent variable regression. Only those variables with significant values can be further interpreted.

Table 30

Regression Analysis of Dependent Variables

Variable	**R^2**	**F**	**Sig.**
Text Recall	0.419	5.040	0.001
Lexical Access	0.129	1.037	0.412
Bridging Inference	0.568	9.201	0.000
Elaborative Inference	0.675	14.553	0.000
Delayed Recall	0.545	8.381	0.000

This table shows that four out of the five regressions were significant and are thus appropriate for further analysis. The lexical access variable did not have a significant regression. This means that the variation in student scores on this variable is not significantly predicted by any of the five reader characteristic variables. This corresponds with the results from the canonical correlation which showed that lexical access had very low redundancy and a small correlation coefficient when compared to the other variables. Lexical access will therefore not be further investigated using the regression analysis.

The four variables with significant regressions (Text Recall, Bridging Inference, Elaborative Inference, and Delayed Recall) were investigated to determine which of the independent variables was significantly predicting the variance in students' scores. This was done by comparing the standardized coefficients of the predictor variables to determine which was having a significant effect. The following tables provide an overview of that investigation in which the significant values are highlighted.

Table 31

Regression Coefficient Between Text Recall and the Independent Variables

Independent Variable	**Standardized Coefficient**	**Sig.**
Working Memory	0.050	0.720
GALT	0.120	0.453
Chemistry Content	0.025	0.918
Path Length Correlation	0.055	0.865
Neighborhood Similarity	0.517	0.032

Table 32

Regression Coefficient Between Bridging Inferences and the Independent Variables

Independent Variable	**Standardized Coefficient**	**Sig.**
Working Memory	0.066	0.589
GALT	0.050	0.717
Chemistry Content	0.193	0.367
Path Length Correlation	0.580	0.042
Neighborhood Similarity	0.023	0.907

Table 33

Regression Coefficient Between Elaborative Inferences and the Independent Variables

Independent Variable	Standardized Coefficient	Sig.
Working Memory	0.143	0.179
GALT	0.029	0.806
Chemistry Content	0.017	0.927
Path Length Correlation	0.874	0.001
Neighborhood Similarity	0.043	0.805

Table 34

Regression Coefficient Between Delayed Recall and the Independent Variables

Independent Variable	Standardized Coefficient	Sig.
Working Memory	0.059	0.635
GALT	0.078	0.583
Chemistry Content	0.080	0.714
Path Length Correlation	0.171	0.549
Neighborhood Similarity	0.491	0.022

These tables show that in each dependent variable only one reader characteristic was significantly predicting the variance in the students' scores. In both Recall variables, Text and Delayed, the Neighborhood Similarity predictor was significant. In both Inference variables, the Path Length Correlation variable was significant. In all cases the significant independent variable was one of the two Pathfinder similarity scores, Path Length Correlations or Neighborhood Similarities. This supports the canonical correlation results that show that the Path Length Correlation score had the highest canonical weight, high redundancy, and a high correlation with the covariate. In the canonical correlation Neighborhood Similarity had the third highest canonical weight of the independent variables, high redundancy and a high correlation with the covariate. Thus it is not surprising that Path

Length Correlation scores explain a large amount of variance in the regression analysis. The results of the regression analysis support the results in the canonical correlation analysis.

Overview of Results

The results of the canonical correlation have shown that there is a significant relationship between the variables involved in this study. This significance is supported by the significance of the canonical correlation function, the magnitude of the function, and the redundancy index of each set of variables. The variables involved in canonical correlation were evaluated using four criteria: (1) the statistical significance of the separate regressions of the independent (reader characteristic) variables; (2) the canonical weights of the variables; (3) the redundancy of the variables; (4) the correlation between each variable and the canonical function (structure correlation). These four analyses were compared to determine which variables play the most important roles in the canonical correlation function. Table 35 summarizes which variables from each set were the top three performing on each analysis.

Table 35

Summary of the Top Three Individual Variable Analyses

	Sig. of Separate Regressions	Canonical Weights	Redundancy	Structure Correlation
Reader Characteristic Set (Independent)[a]	1. PLC	1. PLC	1. PLC	1. PLC
	2. NS	2.WM	2. NS[c]	2. CC
	3. CC	3. NS	2. CC[c]	3. NS
Reading Comprehension Set (Dependent)[a]	See Analysis Below[b]	1. EI	1. EI	1. EI
		2. DR	2. BI	2. BI
		3. LA	3. DR	3. DR

[a] Names of variables have been abbreviated to the following for readability of table:

Path Length Correlation (PLC)
Neighborhood Similarity (NS)
Chemistry Content (CC)
Working Memory (WM)
Elaborative Inferences (EI)
Bridging Inferences (BI)
Delayed Recall (DR)
Lexical Access (LA)

[b]The separate regressions for the reading comprehension variables are not included in this table as they are described in further detail in the next section of this chapter.

[c]Neighborhood Similarity (NS) and Chemistry Content (CC) had equal redundancy values and thus tied for the 2nd slot.

No single analyses in this table can be used to determine which variables are most important to the canonical correlation. It is apparent, however, that some trends do exist. For instance, Path Length Correlation is the top performing independent variable in every analysis. Similarly, Elaborative Inference is the top performing dependent variable in each analysis. Along with Path Length Correlation, the reader characteristic variables that consistently perform well are Neighborhood Similarity and Chemistry Content. With Elaborative Inferences in the reading comprehension variables, Bridging Inference and Delayed Recall are consistently high performing variables. Because they are the top performing variables in this study, these six variables will be the focus of the discussion in Chapter 5.

To further investigate the effect of the reader characteristic variables and to validate the results of the canonical correlation, separate regression analyses on the reading comprehension variables were performed. The results of the separate regressions for each of the reading comprehension variables indicated which reader characteristic variables were significantly contributing to the variance in students' responses. After evaluating each regression's significance, it was found that the Lexical Access regression was not significant. This is interpreted as none of the reader characteristic variables was significantly predicting the variability of the students on this variable, a fact that will be discussed in further detail in the limitations section of Chapter 5.

The regressions for the remaining four variables in the Reading Comprehension set, Text Recall, Bridging Inferences, Elaborative Inferences, and Delayed Recall where all found to be significant. In each regression only one reader characteristic variable was found to significantly predict the reading comprehension variable of interest. Table 36 provides a summary of these results.

Table 36

Summary of Regression Results

Reading Comprehension Variable (Dependent Variable of Regression)	**Reader Characteristic Variable Independent Variable) Found to Be a Significant Predictor**
Text Recall	Neighborhood Similarity
Bridging Inference	Path Length Correlation
Elaborative Inference	Path Length Correlation
Delayed Recall	Neighborhood Similarity

This table highlights the fact that both Recall variables were significantly predicted by the Neighborhood Similarity variable, while both Inference variables were significantly predicted by the Path Length Correlation variable. In all four cases, the reader characteristic variables that significantly predicted the reading comprehension variables were those associated with the Pathfinder program. The implications of the results from both the canonical correlation and the regression analysis will be discussed in Chapter 5.

CHAPTER FIVE

There are many different components of a course that can potentially impact students' understanding of chemistry. Traditionally these components include the students' attendance and participation in lecture, experimentation done in the lab, and efforts completed outside of class including solving problems and reading the text. The purpose of this study was to investigate one of these components of student understanding, namely reading comprehension of a chemistry text. This study determined the nature of relationships among reader characteristics, including schema, logical reasoning, working memory, and factual knowledge on chemistry reading comprehension. Specifically, it investigated the effect the student's reader characteristics had on the different components of reading comprehension. In this study reading comprehension was investigated through the three levels of representation each student creates which were previously discussed in Chapter 2 (Kintsch, 1998). Components of these levels measured in this study included the students' ability to complete the following: encode the textual information (surface level); access relevant lexical information (propositional level); make necessary bridging and elaborative inferences (situational model); and recall general information about the chemistry passages (situational model). These variables were evaluated using a canonical correlation and separate regressions to address the research questions posed in Chapter 1.

This Chapter will discuss the conclusions based on the results presented in Chapter 4. The conclusions will be written in response to the following research questions posed in Chapter 1:

1) What is the nature of the relationships among schema, logical reasoning, factual chemistry knowledge, working memory capacity and reading comprehension of a text when the content is general chemistry?

2) Is the student's schema, logical reasoning, factual chemistry knowledge, or working memory capacity a good predictor of his/her ability to?

a. encodes scientific words in a general chemistry text?

b. access lexical information involved in the understanding of scientific words in a general chemistry text?

c. makes inferences necessary to understand a general chemistry text?

I. Is there a differential effect in bridging vs. elaborative inferences?

d. recall the general ideas presented within a general chemistry text approximately 24 hours after initial reading?

The following Table 37 provides the variable name abbreviation associated with each of the factors described in the research questions.

Table 37

Relationships Between the Factors in the Research Questions and the Variable Name

	Variable Name
Reader Characteristics:	
Schema	Path Length Correlation
	Neighborhood Similarity
Logical reasoning	Logical Reasoning
Factual chemistry knowledge	Chemistry Content
Working memory capacity	Working Memory
Reader Comprehension	
Encoding scientific words	Text Recall
Accessing lexical information	Lexical Access
Making bridging inferences	Bridging Inferences
Making elaborative inferences	Elaborative Inferences
Recall of general ideas.	Delayed Recall

Each of these questions will be addressed individually with a discussion of the conclusions based on the results and where the conclusions fit in the current literature. Limitations of the study will be included along with suggestions for future research in the field of scientific literacy. The chapter will conclude with discussion of the implications of this study for chemical educators.

Research Question 1: What is the nature of the relationships among schema, logical reasoning, factual chemistry knowledge, working memory capacity and reading comprehension of a text when the content is general chemistry?

The results from the canonical correlation analysis indicate that there is a relationship among the variables involved in this study and that this relationship is significant. This was determined by the fact that the canonical function was significant, the magnitude of the function was large, and the redundancy index of each set of variables was acceptable when compared to other studies in relevant fields. In previous studies it has been shown that working memory, logical reasoning, content knowledge, and schema are all accurate predictors of the various aspects of students' reading comprehension skills (Bransford & Johnson, 1972; Cain & Oakhill, 2004; Chomsky, 2006; Ericsson & Kintsch, 1995; Kintsch, 1994; Palladino, Cornoldi, De Beni, & Pazzaglia, 2001; Pichert & Anderson, 1977; van Dijk & Kintsch, 1983). These studies investigated some or all of the reading comprehension variables included in the present study and found various relationships between the different reader characteristic variables and the reading comprehension variables. The significance found in this study shows that the relationship among all the variables involved may be better described when we incorporate them into one cohesive model. The resulting model which combines the variables describes a larger percent of the relationships then any individual variable. As was hypothesized in Chapter 1, a more integrated theory of reading comprehension which includes various reader characteristics and levels of reading comprehension provides a better understanding of students' scientific literacy then the investigation of any one variable alone.

To further investigate each reader characteristic and the reading comprehension variables' roles in the integrated model, the individual variables were evaluated using four analyses which included the following: (1) the statistical significance of the separate regressions of the independent (reader characteristic) variables; (2) the canonical weights of the variables; (3) the redundancy of the variables; (4) the correlation between each variable and the canonical function (structure correlation). The results from these four analyses are discussed here to determine which variables play the largest roles in the integrated model.

The comparison of the four analyses shows that each set of variables (Reader Characteristic and Reading Comprehension), has three variables that are contributing the most to the model (See Chapter 4 for in-depth discussion of analysis). In the Reader Characteristic set these variables are Path Length Correlation, Neighborhood Similarity, and Chemistry Content, with Path Length Correlation. These variables outperform the other two variables on the four analyses described previously. In the Reading Comprehension set these variables are Elaborative Inferences, Bridging Inferences, and Delayed Recall, with Elaborative Inferences outperforming the other two. The implications of these results are discussed here for each set of variables.

Reader Characteristic Set

In the Reader Characteristic set, the variables that contribute the greatest to the canonical correlation function are the Path Length Correlation, Neighborhood Similarity, and Chemistry Content variables. This means that these variables contribute more to the model then the remaining variables (Working Memory and Logical Reasoning). Norris and Phillips (2003) theory of scientific literacy provides an explanation for these results. As stated in Chapter 2, scientific literacy is a complex process that has two major facets, i.e. the *derived* sense of literacy and the *fundamental* sense of literacy (Norris and Phillips, 2003). Being proficient in the *derived* sense of literacy means the student knows the relevant facts in a content area and therefore is able to understand all of the terms and concepts presented in that text. Being proficient in the *fundamental* sense of literacy means the student is able to interpret or critically evaluate the scientific text. Both senses of literacy play important roles in the student's understanding of the text. In this study the *derived* sense of literacy is measured by the Chemistry Content variable and the *fundamental* sense of literacy is measured by the Working Memory, Logical Reasoning, Path Length Correlation, and Neighborhood Similarity variables.

The three reader characteristic variables that contribute the most to the canonical correlation model are Chemistry Content, Path Length Correlation, and Neighborhood Similarity. This means that variables that measured both the *derived* (Chemistry Content) and the *fundamental* (Path Length Correlation and Neighborhood Similarity) senses of literacy are important to the canonical correlation model. This supports the claims of Norris and Phillips (2003) that the factual chemistry knowledge (*derived* sense) is an important but not sufficient component in reading comprehension of scientific text. The fact that the schema variables also contribute significantly shows that being proficient in the chemistry content is not the only contributing factor to reading comprehension. The schema variables account for the second facet of scientific literacy, the *fundamental* sense of literacy. Overall the results from the canonical correlation analysis indicate that both the *derived* sense and the *fundamental* sense of literacy are important factors in students' reading comprehension.

Reading Comprehension Set

While the main focus of the study was to determine how the students' reader characteristics affected their reading comprehension, it is worthwhile to determine which reading comprehension variables account for the greatest variance in the students' reading comprehension responses. In the Reading Comprehension set of variables, the top three performing variables in terms of the amount of variance they describe in the model are Elaborative Inferences, Bridging Inferences, and Delayed Recall. Just and Carpenter (1987), described two stages of reading comprehension. The first stage involved the encoding of the text and understanding of the meaning of the words being read. In this study the Text Recall and Lexical Access variables were used to evaluate this first stage. The second stage is the process by which words are combined to form mental representations of both the relationships of the words found in the text and the relationship between the words in the text and the reader's prior knowledge. In this study Elaborative Inferences, Bridging Inferences, and Delayed Recall were used to evaluate the second stage. Of the five variables included in this set, the three that contributed most to the model, Elaborative Inferences, Bridging Inferences, and Delayed Recall, represent the higher stage of reading comprehension in the Just and Carpenter model. This means that the greatest differences between students' reading of chemistry texts lies not in their ability to encode and understand the words, but rather in their ability to make inferences about the text and integrate that text with their prior knowledge.

As stated in Chapter 2, reading comprehension and the processes involved can be investigated through the three levels of representation of text formed by the student (Kintsch, 1998). These levels include the surface level, propositional level, and situational model. The five variables in this study used to measure the three levels of representation included the following: Text Recall (surface level representation), Lexical Access (propositional level representation), Elaborative Inferences and Bridging Inferences (situational model), and Delayed Recall (situational model). Each subsequent level is more cognitively difficult for the reader to create then the last. In this study, the reading comprehension variables that contributed the most to the study are Elaborative Inferences, Bridging Inferences, and

Delayed Recall. These variables represent the highest level of textual representation (Situational Model). In conclusion both variable sets (Reader Characteristic and Reading Comprehension) the variables contributing the greatest to the model are those that measure higher cognitive processes. These higher processes account for the greatest amount of variance in the student responses.

Previous studies on the effects reader characteristics have on reading comprehension have been limited to one or two variables. This study was conducted to determine if a model that included a larger number of the relevant variables could account for more of the variance in students' responses. The results show that the model in this study including five reader characteristics, (Working Memory, Logical Reasoning, Chemistry Content, Path Length Correlation, and Neighborhood Similarity) account for a larger amount of the variance in students' reading comprehension then previously reported in the literature. The integration of a greater number of variables into one model has resulted in a greater amount of the variance explained. This study has shown that the integrated model of scientific literacy that takes an increased number of variables into consideration provides a better explanation of students' scientific literacy then previous studies which looked at the variables individually.

The results from the canonical correlation have also shown that those reader characteristic variables that are contributing the most (Chemistry Content, Path Length Correlation, and Neighborhood Similarity) represent both the *derived* sense and the *fundamental* sense of scientific literacy. Thus, this study supports the results of past studies in showing that students need both an understanding of the chemistry content (*derived* sense) and the ability to interpret the content (*fundamental* sense) for reading comprehension.

While it was not the focus of this study, the results also indicate which of the reader comprehension variables accounted for the greatest amount of variance in student responses. These variables include both the two inference variables (Bridging and Elaborative) and Delayed Recall. These three variables are those that represent the most difficult level of textual representation for the students to create, namely, the Situational Model. The results of this study show that the variables measuring the top levels of representation are accounting for the greatest differences in students' reading comprehension. In the next research question and sub questions, the individual reading comprehension variables will be further investigated to determine which of the reader characteristic variables has the greatest effect on each reading comprehension variable.

Research Question 2a: Is the student's schema, logical reasoning, factual chemistry knowledge, or working memory capacity a good predictor of his/her ability to encode scientific words in a general chemistry text?

The Text Recall variable was used to investigate the students' ability to encode scientific words. The results from the regression using the Text Recall variable as the dependent variable show that the students' ability to encode scientific words in a general chemistry text is significantly predicted by the reader characteristic variables. When the standardized weights of the reader characteristic variables are compared, it is found that the Neighborhood Similarity variable is the only reader characteristic variable that is significantly predicting the variance in the students' ability to encode the chemistry text. This variable describes 51.7% of the variance in the students' ability to encode chemistry text. As one of the lower levels of textual representation, it is interesting that this variable is not predicted by some of the other reader characteristic variables such as chemistry content or working memory. It is hypothesized that familiarity with the chemistry content or working memory capacity would affect the students' ability to encode the terms but this is not the case in this study. Instead, the students' schema of the chemistry concepts is the most important variable. As the students' schema becomes more structured and closer to that of an expert, the students significantly increase their ability to encode the chemistry text. In the present study this relationship is significant while the other reader characteristic variables' relationships with encoding are not.

Based on these results, the process of encoding information seems to be more complex than originally thought. It involves not only the ability to encode and understand the chemistry terms, but also the integration of these terms with students' prior knowledge or schema. The more expert students' schema, the better students will be able to encode the chemistry text. Thus, the students' reading comprehension may benefit from encountering the information first in class where they are provided with more context and supporting information. By providing them with the content in class prior to reading the text, the instructor could help the students create their schema of the relevant content. Students would then be better prepared to read and encode the information in their chemistry text.

Research Question 2b: Is the student's schema, logical reasoning, factual chemistry knowledge, or working memory capacity a good predictor of his/her ability to access lexical information involved in the understanding of scientific words in a general chemistry text?

As opposed to other studies (Just and Carpenter, 1987) the regression using the Lexical Access variable as the dependent variable was not. This means that none of the reader characteristic variables is significantly predicting the Lexical Access variable. This is in contrast to the results found in studies discussed in Chapter 2. In the present study, the Lexical Access variable is the most difficult of the reading comprehension variables to measure due to the fact that there are no outward behaviors that can be easily measured. This fact was addressed by utilizing a non-traditional eye tracking protocol. The Just and Carpenter (1987) research described in Chapter 2 showed that the length of time on a word reflects the amount of time it takes for a reader to access the meaning of that word. It follows that those students who spent a longer time on the scientific words are taking longer to access the meaning of those words and are therefore less familiar with the terms. Students with a better understanding of the chemistry content or with schemas more like an expert's, should be better able to access the relevant lexical information. The data in this study indicates that this is not the case. Students spent the same amount of time accessing the meaning of the scientific terms regardless of their working memory capacity, logical reasoning ability, knowledge of the chemistry content, or their schema of the relevant information. These results and the reason for their non-significance will be discussed in the limitations section of this chapter.

Research Question 2c: Is the student's schema, logical reasoning, factual chemistry knowledge, or working memory capacity a good predictor of his/her ability to make inferences necessary to understand a general chemistry text?

The Bridging Inference and Elaborative inference variables were used to investigate the students' ability to make inferences. The regressions using the inference variables, Bridging and Elaborative, as the dependent variable were both found to be significant. In both cases the one reader characteristic variable that significantly predicted the students' ability to make inferences was the Path Length Correlation variable. The Path Length Correlation variable represents the correlation of the links in the students' network with the links in the referent expert network. In other words, this is a measure of how the students store the chemistry information in their schema and whether or not their storage is similar to that of experts. The results of this analysis indicate that as the students' storage becomes more like that of experts, their ability to make both bridging and elaborative inferences increases. Without a well structured schema of the relevant content, the students may not be able to make the inferences necessary to understand the material. Bridging and elaborative inferences are necessary to create the third level of textual representation, the situational model (Kintsch, 1998). It is understandable that a well structured schema is needed to successfully create this higher level of representation. As was stated in the encoding section , it may be beneficial to the reading comprehension process to present the chemistry concepts to students prior to asking them to read their chemistry text. This order would provide students with the background information needed to make the inferences necessary to understand the text.

Research Question 2ci: Is there a differential effect in bridging vs. elaborative inferences?

Both types of inferences, bridging and elaborative, were significantly predicted by one of the reader characteristic variables, Path Length Correlation. The amount of variance predicted was, however, different. The Path Length Correlation variable predicted 58% of the variance in students' responses to bridging inference questions, and 87.4% of the variance in students' responses to the elaborative inference questions. This indicates that the quality of the students' schema has a greater effect on their ability to make elaborative inferences then bridging inferences. Elaborative inferences require students to make connections between what they are reading and their prior knowledge while bridging inferences only require students to make connections between concepts within the text. As the students' schema becomes more complete and more similar to that of an expert, they are better able to make the connections between the text they are reading and their prior knowledge. Again, it is important that the prior knowledge be as complete as possible before the student attempts to read the text for understanding. This may require some instruction on the information so that students' schema of the concepts are more complete before they read the text.

Research Question d: Is the student's schema, logical reasoning, factual chemistry knowledge, or working memory capacity a good predictor of his/her ability to recall the general ideas presented within a general chemistry text approximately 24 hours after initial reading?

The Delayed Recall variable was used to investigate the students' ability to recall the general ideas of a chemistry text approximately 24 hours after an initial reading. The regression using Delayed Recall as the dependent variable was found to be significant, with Neighborhood Similarity being the only reader characteristic variable to describe a significant amount of the variance. This means that as students' schema become more expert-like, they are better able to recall the general ideas presented within a general chemistry text after a delay. The Neighborhood Similarity variable accounted for 49% of the variance in the students' responses on the Delayed Recall variable. In this study, the Delayed Recall variable was used to investigate the students' third level of representation—the situational model. Thus, better structured schemas, measured by the Neighborhood Similarity variable, would help students to recall more of the information given in the text. This could be due to the fact that more complete schemas allow students to better integrate the new information with their prior knowledge.

These results provide an important insight to the potential effect of providing the students with conceptual information prior to their reading the chemistry text. If we provide students with the background information of the chemistry concept before we ask them to read the text, we may increase their ability to retain information in the text. Helping students form their schema closer to that of experts prior to reading the concept in the text may prime them to better retain the scientific information.

Overview of Separate Regressions

The four dependent variables with significant regressions were Text Recall (encoding at the surface level), Bridging and Elaborative Inferences (creating the Situational Model), and Delayed Recall (creation and use of the Situational Model). These results support the Canonical Correlation results in which Neighborhood Similarity and Path Length Correlation were two of the top three performing variables. The third, Chemistry Content, was not found to significantly predict the variance in the reading comprehension variables in the regression analysis.

In both of the recall variables (Text Recall and Delayed Recall) the Neighborhood Similarity reader characteristic variable accounted for a significant amount of the variance in students' responses. This indicates that the way students group the different chemistry concepts within a given topic greatly affects their ability to recall the information. Similarly, in both of the inference variables, Bridging Inference and Elaborative Inference, the Path Length Correlation variable accounted for a significant amount of the variance in students' responses. This indicates that the presence and strength of the connections between chemistry concepts within students' schema of a topic greatly affects their ability to make inferences pertaining to that information. Both of these findings support the idea that students' schema of a chemistry topic is an important component in the processes involved in reading comprehension.

In Chapter 1 it was hypothesized that the Pathfinder variables would account for the greatest amount of variance in the processes involved in students' reading comprehension. This hypothesis was made due to the fact that the creation of a complex schema involves more cognitive components then do the other reader characteristic variables. Therefore the reader characteristic variables that measure student's schema, Path Length Correlation and Neighborhood Similarity, would be expected to account for the greatest amount of variance in students' reading comprehension. The results of this study support this hypothesis. The students' schema has been shown to be the most important component in the various processes involved in reading comprehension when the text is a general chemistry text.

Limitations and Future Research

As with any study, limitations may apply to these findings due to the limitations associated with a particular sample and experimental conditions. The content involved in this study is a subset of chemistry topics and therefore the results may not be generalizable to other chemistry concepts and/or scientific fields. It has been proposed, however, that because of the similarities between scientific texts in all fields, success in one field (chemistry) could result in increased success in a student's general scientific literacy (Cromley & Snyder-Hogan, 2010; Klein, 2006; Yore & Treagust, 2006). This means that the results from this study might be used in other fields but replication of the study would first be necessary in those fields.

Another limitation exists due to the fact that the Working Memory regression in the Canonical Correlation analysis was not significant and therefore did not account for a significant amount of the variance in the Reading Comprehension set of variables. This result is in contrast to those studies involving working memory, discussed in Chapter 2 (Atkinson and Shiffrin, 1968; Cain and Oakhill, 2004; Miller, 1956; Newell and Simon, 1972). In the literature working memory affected various cognitive processes including propositional structures, reading goals, lexical access, concept formation and access of schema. So why then, did the Working Memory variable in this study not describe a significant amount of the variance in the students' reading comprehension responses? It may be, in part, due to the limited range of responses on this variable. Table 38 (Chapter 4) presented here, presents the range, mean, and standard deviation of the Working Memory variable.

Table 38

Descriptive Statistics of the Working Memory Variables

Variable	**Range**	**Mean**	**Standard Dev.**
Working Memory[a]	5—9	7.16	0.90

[a]Working Memory scores can range from 0 to the highest number of digits the person can remember.

This shows that all students scored between five and nine on the forward digit span test, this is a small range. In the studies reported in Chapter 2, the range of working memory scores were much larger. Table 39 provides an overview of the ranges found in the studies reported.

Table 39

Working Memory Score Ranges

Reference	Cain & Oakhill, 2004	Atkinson and Shiffrin, 1968	Miller, 1956	Newell and Simon, 1972
Working Memory Range	2 to 21	3 to 14	2 to 14	1 to 15

While it was possible during the present study for students to obtain any of these scores on the forward digit span test, this group of students scored in the five to nine range. It may be that a larger number of students or a targeted pool of students who had a wider range of Working Memory scores, would result in a significant regression of the Working Memory variable. Further studies should be conducted to investigate this possibility.

One final limitation to this study came from the investigation of the Lexical Access regression. This regression indicates that none of the reader characteristic variables significantly predicts students' scores on the Lexical Access reading comprehension variable. This is in direct contrast to the theories presented in Chapter 2, which state that the reader characteristics chosen for the this study should impact reading comprehension (Hoffman & Subramaniam, 1995; Just and Carpenter, 1980; Reichle et al., 2006). In the studies by Just and Carpenter (1980) the researchers investigated the difference between readers who were familiar with the text and those who were unfamiliar with the text. The results in this study were based upon participants with a wide range of expertise. It is possible that the range of expertise in the present study did not have a wide enough range of chemistry ability to show an effect on the students' lexical accessing ability. To investigate this further, differences in the science background of the students and its effect on lexical access were evaluated. No significant difference was found in the amount of time spent reading the chemistry passages for students of this study's different science backgrounds. The same was true for the *percentage* of time spent reading scientific words. This can be interpreted as students who are chemistry majors or had completed a greater number of chemistry courses did not read the paragraphs or access the meanings of the scientific words faster than those students who did not have as strong a chemistry background. This leads the researcher to believe that in order to find differences in the process of lexical access, participants with a greater difference in expertise would be necessary. Further studies using participants with a larger range of chemical expertise are suggested to determine whether the non-significant finding of the lexical access in this study is valid.

Implications for Teaching

An important goal of educational research is to better understand the process of student learning of a given subject. To address this goal, the research must often investigate only one component of the learning process. This study investigated reading comprehension which is one of the many components that can potentially impact students' understanding of chemistry. By investigating reading comprehension, we may gain a better understanding of how to increase the students' ability to understand chemistry texts. This may in turn result in an increase in students' understanding of chemistry content.

In a general chemistry classroom the traditional use of the textbook is to assign students reading prior to covering the material in lecture. This often leads to frustration on the part of instructors when their students don't understand the material even after reading the text prior to class. Instructors may believe that students do not spend an adequate amount of time reading the text. The results of this study indicate that this assumption may be incorrect. It may be that the students are indeed reading the text but are having a difficult time understanding the information found there. The students' organization of the material in their minds (schema) is important to their ability to understand the text. When we consider the processes involved in the students' reading comprehension, it appears that it is beneficial for the students to encounter the content in class first, and read the text second. This way their minds have been primed to better understand the concepts and retain the information for a longer period of time. With a more complete schema the students will be better able to encode the textual information, make the inferences necessary to read the text, and retain the information after it has been read. Many have had the experience of reading something not fully understood at the time and then later having the content explained in such a way that the text makes more sense. Giving students the background information prior to having them read the text may result in better understanding of the information and thus more effective incorporation of that information into their schema. This relatively small change in teaching method could have a large impact on student understanding of chemistry.

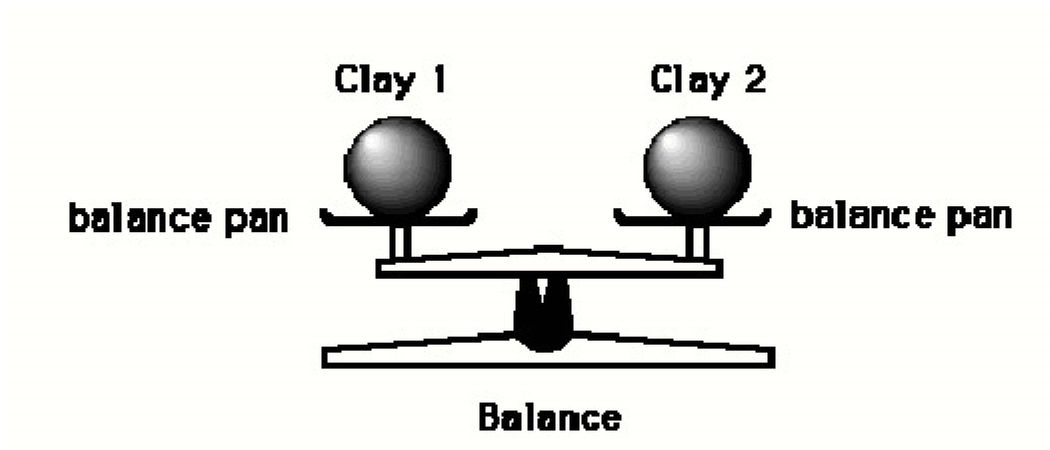
Appendix A

Group Assessment of Logical Thinking (GALT)

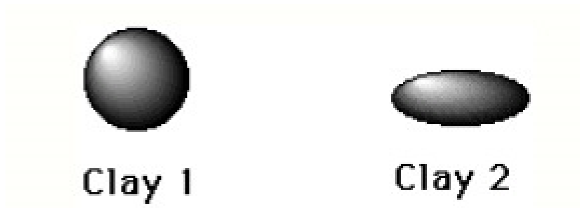
Question 1:Piece of Clay

Tom has two balls of clay. They are the same size and shape.

When he places them on the balance, they weigh the same.



The balls of clay are removed from the balance pans. Clay 2 is flattened like a pancake.



WHICH OF THESE STATEMENTS IS TRUE?

Answer

a. The pancake-shaped clay weighs more.
✓ b. The two pieces weigh the same.
c. The ball weighs more.

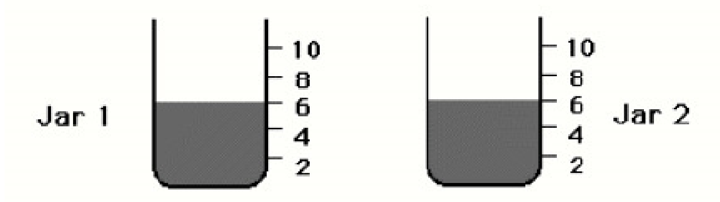
SELECT THE REASON FOR YOUR ANSWER:

✓ d. You did not add or take away any clay.
e. When clay 2 was flattened like a pancake, it had greater area.
f. When something is flattened, it loses weight.
g. Because of its density, the round ball had more clay in it.

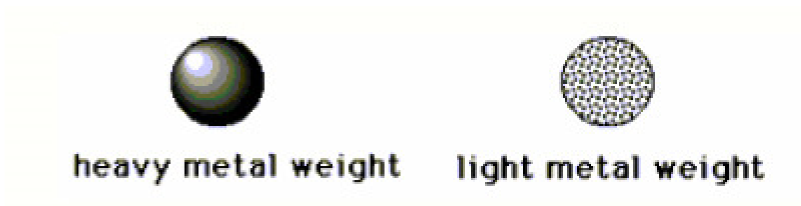
Question 2: Metal Weights

Linn has two jars. They are the same size and shape.

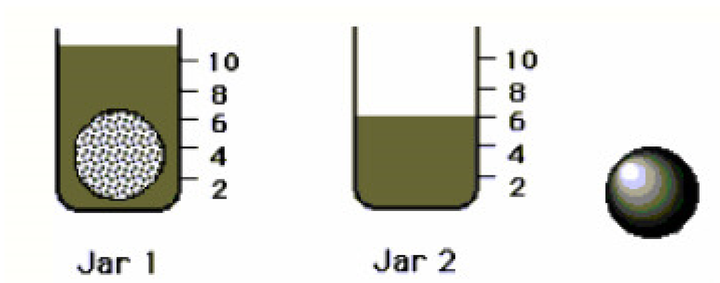
Each is filled with the same amount of water.



She also has two metal weights of the same volume. One weight is light. The other weight is heavy.



She lowers the light weight into jar 1. The water level in the jar rises and looks like this:



IF THE HEAVY WEIGHT IS LOWERED INTO JAR 2, WHAT WILL HAPPEN?

Answer

a. The water will rise to a higher level than in jar 1.

b. The water will rise to a lower level than in jar 1.

✓ c. The water will rise to the same level as in jar 1.

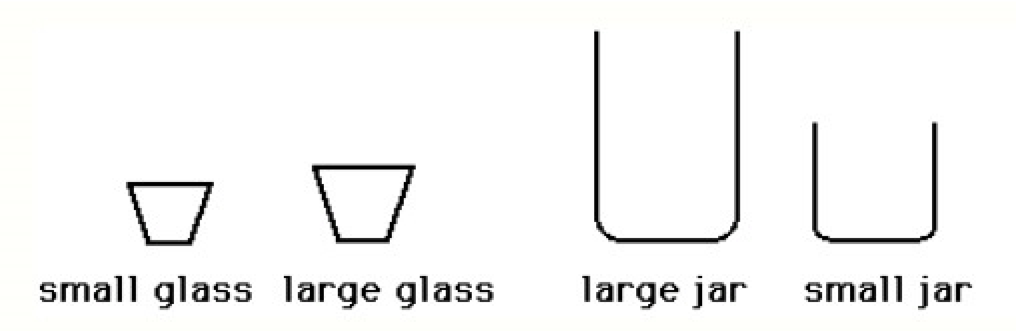
SELECT THE <u>REASON</u> FOR YOUR ANSWER:

✓ d. The weights are the same size so they will take up equal amounts of space.

e. The heavier the metal weight, the higher the water will rise.
f. The heavy metal weight has more pressure, therefore the water will rise.
g. The heavier the metal weight, the lesser the water will rise.

Question 3:Glass Size #2

The drawing shows two glasses, a small one and a large one. It also shows two jars, a small one and a large one.



It takes 15 small glasses of water or 9 large glasses of water to fill the large jar. It takes 10 small glasses of water to fill the small jar.

HOW MANY LARGE GLASSES DOES IT TAKE TO FILL THE SAME SMALL JAR?
Answer

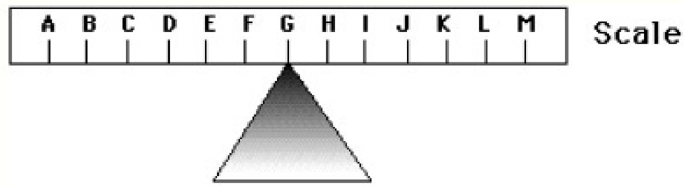
a. 4
b. 5
✓ c. 6
d. other

SELECT THE <u>REASON</u> FOR YOUR ANSWER:

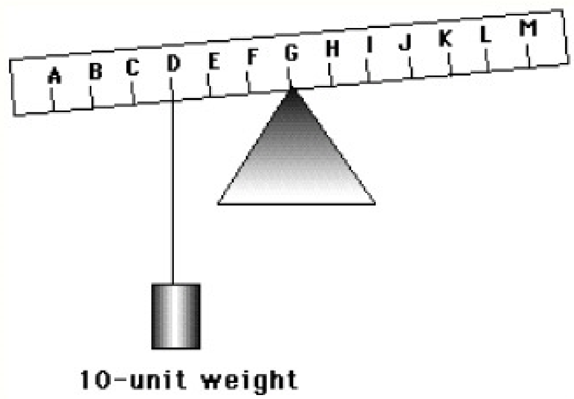
e. It takes five less small glasses of water to fill the small jar. So it will take five less large glasses of water to fill the same jar.
✓ f. The ratio of small to large will always be 5 to 3.
g. The small glass is half the size of the large glass. So it will take about half the number of small glasses to fill up the same small jar.
h. There is no way of predicting.

Question 4: Scale #1

Joe has a scale like the one below.



When he hangs a 10-unit weight at point D, the scale looks like this:



WHERE WOULD HE HANG A 5-UNIT WEIGHT TO MAKE THE SCALE BALANCE AGAIN?

Answer

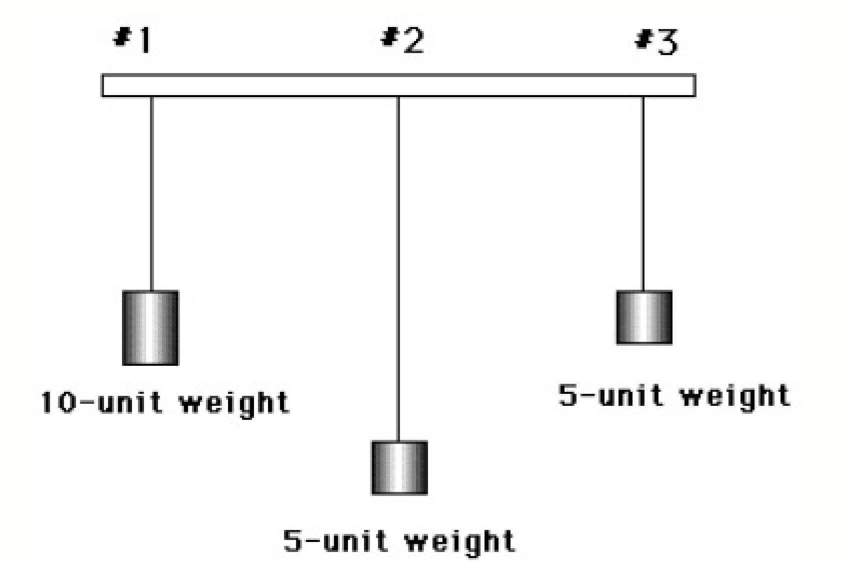
a. at point J
b. between K and L
c. at point L
d. between L and M
✓ e. at point M

SELECT THE <u>REASON</u> FOR YOUR ANSWER:

✓ f. It is half the weight so it should be put at twice the distance.
g. The same distance as the 10-unit weight, but in the opposite direction.
h. Hang the 5-unit weight further out, to make up for it being smaller.

Question 5:Pendulum Length

Three strings are hung from a bar. String #1 and #3 are of equal length. String #2 is longer. Charlie attaches a 5-unit weight at the end of string #2 and at the end of string #3. A 10-unit weight is attached at the end of string #1. Each string with a weight can be swung.



Charlie wants to find out if the length of the string has an effect on the amount of time it takes the string to swing back and forth.
WHICH STRING AND WEIGHT WOULD HE USE FOR HIS EXPERIMENT?

Answer

string #1 and #2

string #1 and #3

✓ string #2 and #3

string #1, #2 and #3

string #2 only

SELECT THE REASON FOR YOUR ANSWER:

The length of the strings should be the same. The weights should be different.

Different lengths with different weights should be tested.

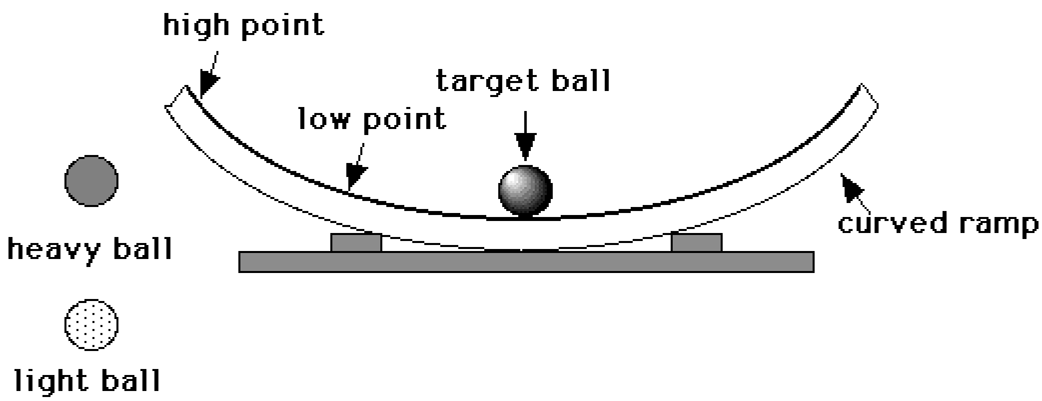
All strings and their weights should be tested against all others.

Only the longest string should be tested. The experiment is concerned with length not weight.

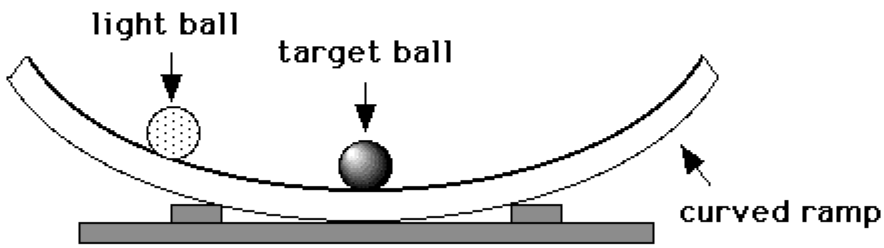
✓ Everything needs to be the same except the length so you can tell if length makes a difference.

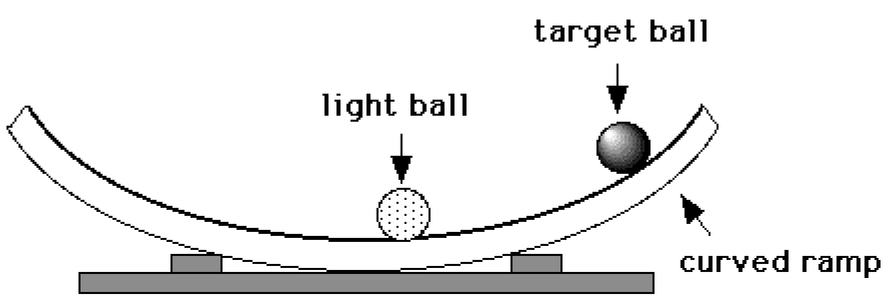
Question 6: Ball #1

Eddie has a curved ramp. At the bottom of the ramp, there is one ball called the target ball.



Eddie released the light ball from the low point. It rolled down the ramp. It hit and pushed the target ball up the other side of the ramp.





He wants to find out if the point a ball is released from makes a difference in how far the target ball goes.

TO TEST THIS WHICH BALL WOULD HE NOW RELEASE FROM THE HIGH POINT?

Answer the heavy ball

✓ the light ball

SELECT THE REASON FOR YOUR ANSWER:

He started with the light ball, he should finish with it.

He used the light ball the first time. The next time he should use the heavy ball.

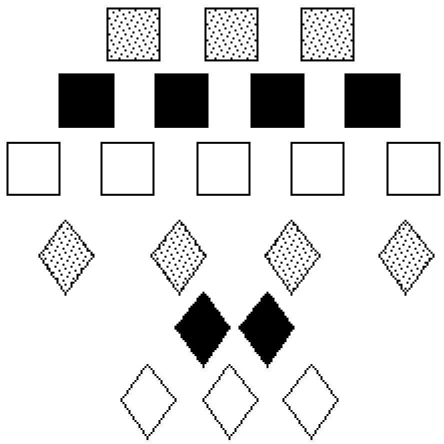
The heavy ball would have more force to hit the target farther.

✓ The light ball would have to be released from the high point in order to make a fair comparison.

The same ball must be used as the weight of the ball does not count.

Question 7: Squares and Diamonds

In a cloth sack, there are



All of the square pieces are the same size and shape. The diamond pieces are also the same size and shape. One piece is pulled out of the sack.

WHAT ARE THE CHANCES THAT IT IS A SPOTTED PIECE?

Answer

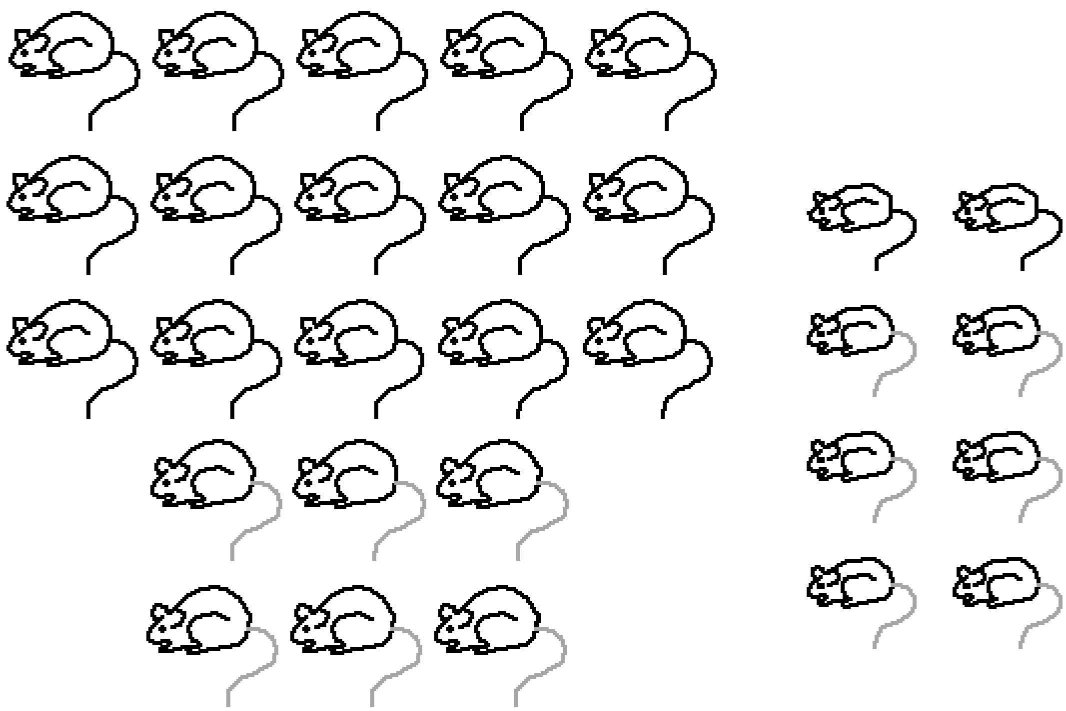
✓ 1 out of 3
1 out of 4
1 out of 7
1 out of 21
other

SELECT THE REASON FOR YOUR ANSWER:

There are twenty-one pieces in the cloth sack. One spotted piece must be chosen from these.
One spotted piece needs to be selected from a total of seven spotted pieces.
✓ Seven of the twenty-one pieces are spotted pieces.
There are three sets in the cloth sack. One of them is spotted.
1/4 of the square pieces and 4/9 of the diamond pieces are spotted

Question 8: The Mice
A farmer observed the mice that live in his field. He found that the mice were either fat or thin. Also, the mice had either black tails or white tails.

This made him wonder if there might be a relation between the size of a mouse and the color of its tail. So he decided to capture all of the mice in one part of his field and observe them. The mice that he captured are shown below.



DO YOU THINK THERE IS A RELATION BETWEEN THE SIZE OF THE MICE AND THE COLOR OF THEIR TAILS (THAT IS, IS ONE SIZE OF MOUSE MORE LIKELY TO HAVE A CERTAIN COLOR TAIL AND VICE VERSA)?

Answer

✓ Yes

No

SELECT THE REASON FOR YOUR ANSWER:

✓ 5/7 of the fat mice have black tails and 3/4 of the thin mice have white tails.

Fat and thin mice can have either a white tail or a black tail.

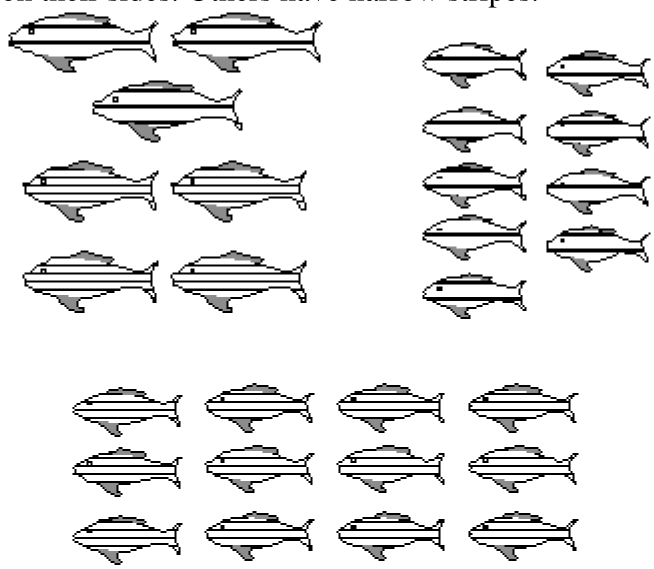
Not all fat mice have black tails. Not all thin mice have white tails.

17 mice have black tails and 12 have white tails.

21 mice are fat and 8 mice are thin.

Question 10: The Fish

Some of the fish below are big and some are small. Also some of the fish have wide stripes on their sides. Others have narrow stripes.



S THERE A RELATIONSHIP BETWEEN THE SIZE OF THE FISH AND THE KIND OF STRIPES IT HAS (THAT IS, IS ONE SIZE OF FISH MORE LIKELY TO HAVE A CERTAIN TYPE OF STRIPES AND VICE VERSA)?

Answer

Yes

✓ No

SELECT THE REASON FOR YOUR ANSWER:

Big fish and small fish can have either wide or narrow stripes.

✓ 3/7 of the big fish and 9/21 of the small fish have wide stripes.

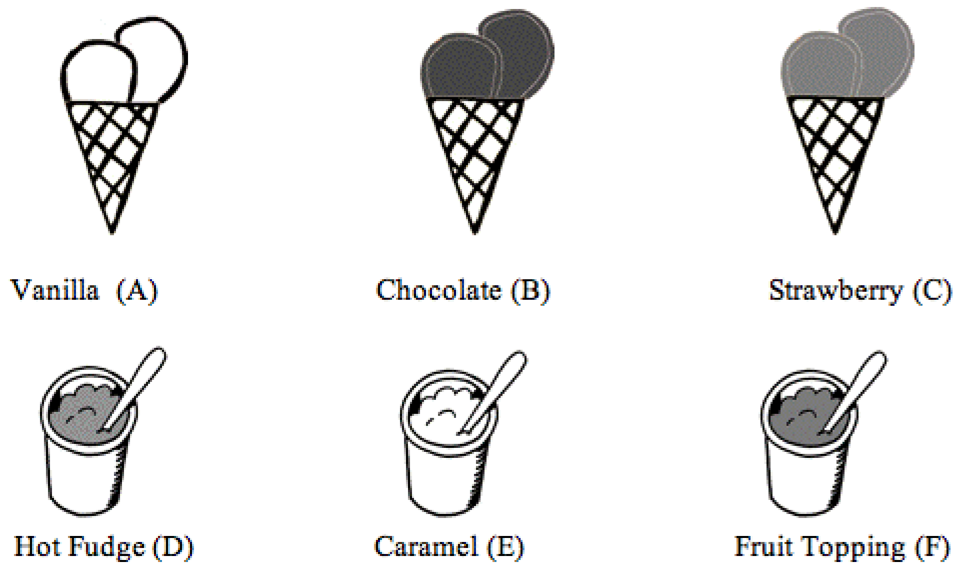
7 of the fish are big and 21 are small.

Not all big fish have wide stripes and not all small fish have narrow stripes.

12/28 of fish have wide stripes and 16/28 of fish have narrow stripes.

Question 11: Ice Cream

After dinner, some students decide to go out for ice cream. The shop offers three flavors of ice cream: Vanilla (A), Chocolate (B), and Strawberry (C), and three toppings: Hot fudge (D), Caramel (E), and Fruit topping (F).



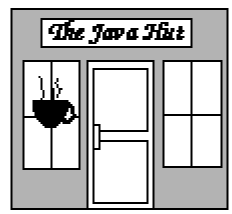
One possible sundae is A-D, (vanilla ice cream and hot fudge). LIST ALL OTHER POSSIBLE SUNDAES, CONSISTING OF ONE ICE CREAM FLAVOR AND ONE TOPPING EACH.

Question 12: Shopping Center

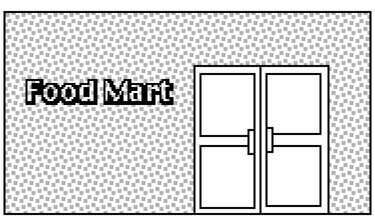
In a new shopping center, 4 stores are going to be placed on the ground floor. A BARBER SHOP (B), a DISCOUNT STORE (D), a GROCERY STORE (G), and a COFFEE SHOP (C) want to locate there.

Barber Shop (B)

Discount Store (D)



Coffee Shop (C)



Grocery Store (G)

One possible way that the stores could be arranged in the 4 locations is BDGC. Which means the BARBER SHOP first, the DISCOUNT STORE next, then the GROCERY STORE and the COFFEE SHOP last.

LIST ALL THE POSSIBLE WAYS THAT THE STORES CAN BE LINED UP IN THE FOUR LOCATIONS.

Appendix B

Multiple Choice Chemistry Content Test

Please complete the following fourteen question chemistry quiz by circling your answer. Only select one answer per question.

1. Which statement concerning the structure of the atom is correct?
    a. Protons and neutrons have most of the mass and occupy most of the volume of the atom.
    b. Electrons have most of the mass and occupy most of the volume of the atom
    c. Electrons have most of the mass but occupy very little of the volume of the atom.
    d. Protons and neutrons have most of the mass but occupy very little of the volume of the atom.

2. How many grams of aluminum chloride can one obtain from 6.00 mol of barium chloride?
    i. $Al_2(SO_4)_3 + 3BaCl_2 \rightarrow 3BaSO_4 + 2AlCl_3$

Atomic Molar Masses	
Al	27.0 g·mol^{-1}
Cl	35.5 g·mol^{-1}

    b. 1250 g
    c. 801 g
    d. 534 g
    e. 134 g

3. In all neutral atoms, there are equal numbers of
    a. Protons and neutrons
    b. Positrons and electrons
    c. Neutrons and electrons
    d. Electrons and protons
4. A sodium ion differs from a sodium atom in that the sodium ion
    a. Is more reactive.
    b. Has fewer electrons.
    c. Is an isotope of sodium.
    d. Exists only in solution.
    e. Has a negative charge on its nucleus.
5. Isotopes of an element differ in

a. Number of electrons in the outermost shell.
b. Number of neutrons in the nucleus.
c. Combining number.
d. Number of protons in the nucleus.
e. Atomic number.

6. One of the tin isotopes has 50 protons and 63 neutrons. Another isotope of tin might have
   a. 50 protons and 0 neutrons.
   b. 50 protons and 62 neutrons.
   c. 49 protons and 63 neutrons.
   d. 63 protons and 63 neutrons.
   e. 63 protons and 50 neutrons.
7. When the equation ? Sb + ? $Cl_2 \rightarrow$? $SbCl_3$ is correctly balanced, the sum of the coefficients is
   a. 2
   b. 3
   c. 6
   d. 7
   e. 9
8. The number of electrons in an atom of any free element is equal to
   a. The number of protons in the nucleus.
   b. The number of neutrons in the nucleus.
   c. The sum of the number of neutrons and protons.
   d. The atomic mass less the number of protons.
   e. The atomic number less the number of neutrons.
9. All positive ions differ from their corresponding atoms by having
   a. Larger diameters.
   b. Fewer electrons.
   c. A charge of +1.
   d. Greater atomic masses.
   e. Stronger metallic properties.
10. To what mass does the expression "one mole of copper (II) sulfide" refer?
    a. The mass of copper (II) sulfide needed to occupy one liter.
    b. The mass of solid copper (II) sulfide needed to occupy 22.4 L at STP.
    c. The mass of copper combined with sulfur in one molecule.
    d. The mass of copper (II) sulfide in grams equal to one formula molar mass.

e. The mass in grams of one molecule of copper (II) sulfide.

11. The elements in an ionic compound are held together by
    - **a. Electrostatic forces of attraction.**
    - b. Van der Waals forces.
    - c. The spin of paired electrons.
    - d. The formation of hybrid orbitals.
    - e. An electron pair.

12. What type of compound is formed when two elements with greatly different electronegativities (positive and negative) combine?
    - **a. Ionic.**
    - b. Molecular.
    - c. Covalent.
    - d. Network covalent.
13. Which substance exhibits ionic bonding?
    - a. H_2O
    - b. CO_2
    - c. SiC
    - **d. NaBr**
14. The total number of atoms represented by $C_3H_5(OH)_3$?
    - a. 14
    - **b. 8**
    - c. 6
    - d. 5
15. Which substance has bonds with greatest ionic character?
    - a. F_2
    - b. OF_2
    - **c. BaF_2**
    - d. CuF_2
16. In the reaction $4Al + 3O_2 \rightarrow 2Al_2O_3$ how many moles of aluminum oxide, Al_2O_3, are produced from one mole of aluminum, Al?
    - **a. 0.5**
    - b. 2
    - c. 3
    - d. 4
17. Consider the unbalanced expression ? $CH_3CH_2CHO_{(l)}$ + ? $O_{2(g)} \rightarrow$? $CO_{2(g)}$ + ? $H_2O_{(g)}$ Which set of coefficients balances the equation?

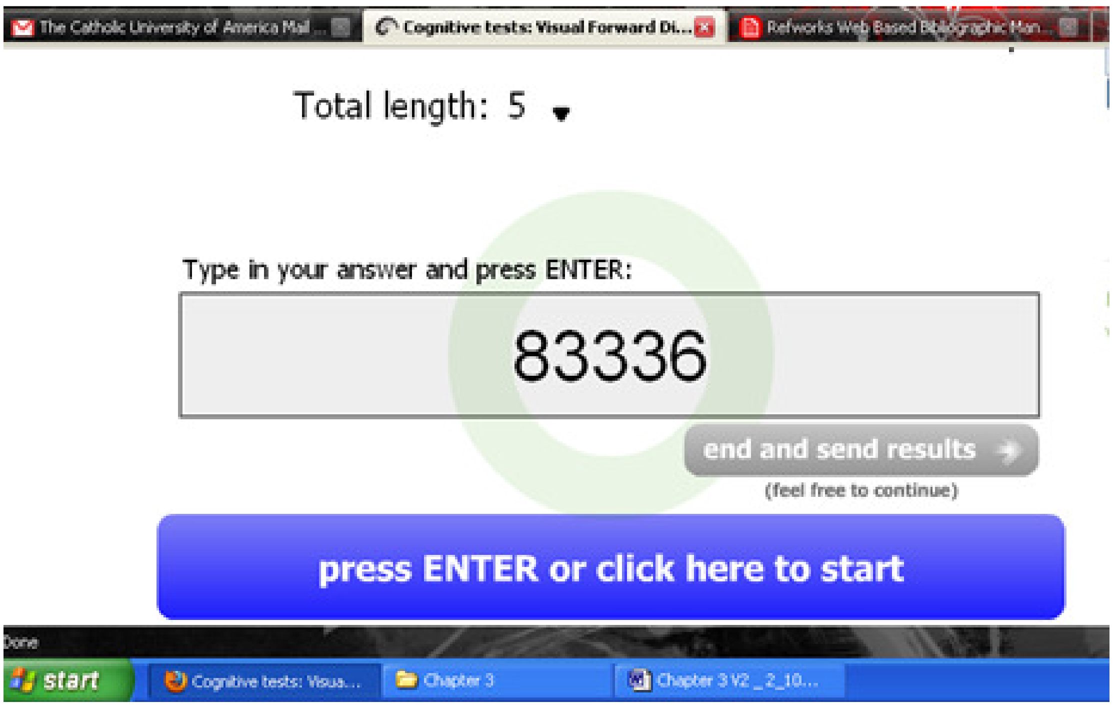
a. 2, 8, 3, 6
b. 3, 8, 6, 6
c. 1, 4, 3, 2
d. 1, 8, 3, 3
e. 1, 4, 3, 3

18. C_3H_8 + ? $O_2 \rightarrow 3CO_2 + 4H_2O$ How many moles of oxygen are needed for the complete combustion of one mole of propane, C_3H_8?
    a. 10
    b. 2.0
    c. 3.0
    d. 5.0

19. What is the total number of atoms represented by $5Al(C_2H_3O_2)_3$ is
    a. 22
    b. 60
    c. 71
    d. 84
    e. 110

Appendix C

Digit Span Test for Working Memory



Appendix D

Key Terms used by Pathfinder Program

<u>Atoms, Ions and Molecules</u>
Atom
Electron
Nucleus
Proton
Neutron
Isotope
Ion
Cation
Anion
Charge
Molecule
Polyatomic Ion
Ionic Compound
Positive
Negative
Molecular Ion

<u>Chemical Bonding</u>
Chemical Bond
Ionic Bond
Covalent Bond
Octet Rule
Attract
Share
Electron Configuration
Lewis Structure
Bond Polarity
Nonpolar Covalent Bond
Polar Covalent Bond
Electronegativity
Dipole
Resonance Structure
Single Bond
Double Bond

Stoichiometry

Stoichiometry
Chemical Equation
Reactant
Product
Subscript
Balanced
Unbalanced
Mole
Avogadro's Number
Molar Mass
Conversion
Limiting
Excess
Theoretical Yield
Actual Yield
Percent Yield

Thermodynamics

Thermodynamic Energy
Thermochemistry
Internal Energy
Kinetic Energy
Potential Energy
Chemical Energy
Thermal Energy
Joule
Calorie
System
Surroundings
Work
Heat
Energy
Endothermic
Exothermic

Appendix E

General Chemistry Passages and Instructions

Instructions

1. You will now be shown a chemistry passage. Please read the passage until you understand the passage to the best of your ability.
2. After you finish reading the text and feel as though you understand the information click the → button to progress to the next page.
3. On this page you will find a space for you to type. In this space type as much of the text as you can recall from the previous passage.
4. Once you have finished typing as much of the passage as you can remember press the → key to progress to the next page.
5. On the following pages you will be asked six multiple choice questions based on the chemistry passage you just read. Once you have completed these questions press the → key to progress to the next chemistry passage.
6. You will repeat this entire process for a total of 3 chemistry passages on this topic.

Page Break

If at any time you would like to stop this process tell the researcher and the procedure will be ended. If you understand these instructions please press the → key to progress to an example.

Page Break

<u>The Tooth</u>
Kristen's tooth had been hurting her all day. Her mother told her that she should make an appointment for later that afternoon. Kristen agreed with her mother so she made a phone call and later that day found herself sitting in a chair. The tooth was worse than Kristen thought and had to be pulled. Fortunately for Kristen, the procedure was painless.

Passages

<u>Atoms, Ions, and Molecules</u>

1. *The Formation of Ions*
The nucleus of an atom is unchanged by ordinary chemical processes, but atoms can readily gain or lose electrons. Electrons can be removed from or added to a neutral atom. The charged particle is called an ion. An ion with a positive charge is called a cation; a negatively charged ion is called an anion. The sodium atom for example, which has 11

protons and 11 electrons, easily loses one electron. The cation has 11 protons and 10 electrons, so it has a net charge of 1+. The net charge on an ion is represented by a superscript. The superscripts +, 2+, and 3+ indicate a net charge resulting from the loss of one, two, or three electrons, respectively; -, 2-, and 3- result from the gain of one, two, or three electrons respectively.

2. *Protons, electrons, neutrons, and isotopes*
What makes an atom of one element different from an atom of another element? All atoms of an element have the same number of protons in the nucleus. This number is different for different elements. Furthermore, because an atom has no net electrical charge, the number of electrons in it must equal its number of protons. Carbon has six neutrons, although some atoms have more and some have less. Atoms of a given element that differ in the number of neutrons and consequently in mass, are called isotopes.

3. *The formation of an ionic compound*
A great deal of chemical activity involves the transfer of electrons between substances. Ions form when one or more electrons transfer from one neutral atom to another. When elemental sodium is allowed to react with elemental chlorine, an electron transfers from a neutral sodium atom to a neutral chlorine atom. We are left with the Na^+ and the Cl^- which bind together through electrostatic forces to form the compound NaCl, which we know better as common table salt. Sodium chloride is an example of an ionic compound, a compound that contains both positively and negatively charged ions.

<u>Stoichiometry</u>

1. *How coefficients and subscripts are used for balancing equations*
In balancing equations, it is important to understand the difference between a coefficient in front of a formula and a subscript in a formula. Notice that changing a subscript in a formula – from H_2O to H_2O_2, for example—changes the identity of the chemical. The substance hydrogen peroxide, is quite different from water. Subscripts should never be changed in balancing an equation. In contrast, placing a coefficient in front of a formula changes only the amount and not the identity of the substance. Thus, $2H_2O$ means two molecules of water, $3H_2O$ means three molecules of water, and so forth.

2. *The mole*
The mole is the SI unit for the amount of a substance. A mol is defined as the amount of a substance that contains the same number of entities as there are atoms in exactly 12 g of carbon-12. This number is called Avogadro's number, in honor of the 19th-century Italian physicist Amedeo Avogadro, and as you can tell from the definition it is enormous. One mol contains 6.022×10^{23} entities (to four significant figures). However, the mole is not just a counting unit, like the dozen, which specifies the number of objects. The definition of the mole specifies the number of objects in a fixed mass of substance. Therefore, 1 mole of a substance represents a fixed number of chemical entities and has a fixed mass.

3. *Theoretical yield, actual yield, and percent yield*

The quantity of product that is calculated to form when all of the limiting reactant reacts is called the theoretical yield. The amount of product actually obtained in a reaction is called the actual yield. The actual yield is almost always less than (and can never be greater than) the theoretical yield. There are many reasons for this difference. Part of the reactants may not react, for example, or they may react in a way different from that desired (side reactions). In addition, it is not always possible to recover all the reaction product from the reaction mixture. The percent yield of a reaction relates the actual yield to the calculated yield.

Appendix F

Bridging and Elaborative Inference Questions

Atoms, Ions, and Molecules

1. *The Formation of Ions*
Bridging Inference Qs
1. Removing electrons from or adding electrons to a neutral atom form a
    a. Nucleus
    b. **Charged particle**
    c. Base
    d. Acid
2. If a sodium atom looses one electron it becomes
    a. An anion
    b. A neutral atom
    c. An isotope
    d**. A cation**
3. The superscripts -, 2-, and 3- represent
    a. Coefficients
    b. **A net charge**
    c. The number of atoms in a molecule
    d. The number of ions

Elaborative Inference Qs
1. A neutral atom contains the following
    a. A nucleus and ions
    b**. A nucleus and electrons**
    c. Only a nucleus
    d. Only electrons
2. If a sodium atom has 11 protons and 11 electrons
    a. **It is a neutral atom**
    b. It is an anion
    c. It is a cation
    d. It is an isotope
3. The net charge of an ion can be found by
    a. Counting the number of electrons
    b. **Finding the difference between the number of electrons and protons**
    c. Finding the difference between the number of electrons and neutrons
    d. Counting the number of protons in a neutral atom

2. *Protons, electrons, neutrons, and isotopes*

Bridging Inference Q's

1. The number of ______ is different for different elements.
    a. Electrons
    b. Neutrons
    c. Nuclei
    d. Protons
2. A Carbon _______ has six neutrons.
    a. Proton
    b. Neutron
    c. Atom
    d. Compound
3. By changing _________, you can have an effect on an atom's mass but still retain the element's identity.
    a. Protons
    b. Neutrons
    c. Electrons
    d. Nuclei

Elaborative Inference Qs

1. The basic unit of an element is:
    a. A molecule.
    b. A compound.
    c. Charged.
    d. An atom.
2. An atom consists of the following
    a. A nucleus, protons, neutrons, and electrons.
    b. Protons, neutrons, and electrons.
    c. Neutrons, isotopes, protons, and electrons.
    d. Protons, neutrons, and a nucleus.
3. The atoms of a given element can differ in their number of
    a. Protons and electrons
    b. Electrons and Neutrons
    c. Protons and Isotopes
    d. Electrons

3. *The formation of an ionic compound*

Bridging Inference Qs

1. The formation of ions is a form of
    a. **Chemical activity**
    b. Balancing chemical equations.
    c. Electrostatic forces
    d. Isotope formation
2. The term 'elemental sodium' refers to
    a. A sodium cation
    b. A sodium anion
    c. A neutral sodium atom
    d. A reactive sodium atom
3. Sodium chloride is represented as
    a. Na^+ and Cl^-
    b. SoCl
    c. NaCl
    d. Na and Cl
    e.

Elaborative Inference Qs

1. A sodium atom is represented by the following symbol
    a. Cl
    b. Na^+
    c. Na
    d. So
2. Na^+ and Cl^- are
    a. Neutral elements
    b. Neutral atoms
    c. Isotopes
    d. Ions
3. A compound is
    a. Two or more substances combined together through chemical activity.
    b. Two atoms combined together through chemical activity.
    c. Two atoms that share electrons.
    d. Two or more elemental atoms.

Stoichiometry

1. *How coefficients and subscripts are used for balancing equations*
Bridging Inference Qs

1. Which of the following is the symbolic representation for water?
    a. H2O
    b. H^2O
    c. H_2O
    d. H_2O_2
2. Which of the following is a chemical
    a. H_2O
    b. H
    c. H_2O_2
    d. All of the above
3. Which of the following represents the substance hydrogen peroxide
    a. H_2O
    b. $2H_2O$
    c. $3H_2O$
    d. H_2O_2

Elaborative Inference Qs

1. In chemistry, an equation is
    a. Symbolic representation of a chemical reaction where the products are on the left side and the reactants are on the right side.
    b. The steps necessary to determine the mathematical proportions of products.
    c. Symbolic representation of a chemical reaction where the reactants are on the left side and the products are on the right side.
    d. A representation of the equilibrium where the products equal the reactants.
2. In which of the following symbolic representations is the subscript highlighted?
    a. CO^-
    b. $2SO_4$
    c. $2SO_4$
    d. NaCl
3. A molecule refers to
    a. A group of at least two atoms held together by bonds, which is considered as having its own identity.
    b. A combination of two or more atoms.
    c. A single atom of any element.
    d. The simplest form of mater.

2. *The mole*

Bridging Inference Qs

1. The abbreviation mol stands for
   a. A molecule.
   b. A Mole.
   c. Avogadro's number.
   d. Mass.
2. Which of the following amounts refers to Avogadro's number?
   a. The amount of moles in a dozen.
   b. The moles in an atom.
   c. The number of atoms in exactly 12 g of carbon-12.
   d. The number of moles in one atom of carbon-12.
3. Which of the following numbers is Avogadro's number written with four significant figures?
   a. 6.022
   b. $x10^{23}$
   c. 4.0×10^{23}
   d. $6.022x10^{23}$

Elaborative Inference Qs

4. Using significant figures refers to
   a. Each of the digits of a number that are used to express it to the required degree of accuracy, starting from the first nonzero digit.
   b. Using only those amounts that are important to the calculation of interest.
   c. Using only positive numbers in a chemical calculation.
   d. Reporting only changes in amounts that are statistically significant.
5. A SI unit is
   a. A mole.
   b. A standard unit determined by the American Chemistry Society.
   c. A standard unit determined by the International System of Units.
   d. A meter.
6. Carbon-12 is a carbon atom
   a. That weighs 12 grams.
   b. That contains 6 protons, 6 neutrons, and 6 electrons.
   c. That contains 12 protons, 12 neutrons, and 12 electrons.
   d. That weighs the same as 1 mole.

3. *Theoretical yield, actual yield, and percent yield*

Elaborative Inference Qs

1. The calculated yield can also be referred to as
    a. Percent yield
    b. Actual yield
    c. Theoretical yield
    d. None of the above
2. When comparing the actual yield to the theoretical yield
    a. There will always be a difference between the two.
    b. There will never be a difference between the two.
    c. There may or may not be a difference between the two.
    d. The theoretical yield will always be less than the actual yield.
3. The reaction mixture includes
    a. Only the reactants.
    b. Only the products.
    c. The reactants and the products.
    d. The products in solution.

Elaborative Inference Qs

4. The product in a chemical reaction is
    a. The substances combined to start the chemical reaction.
    b. The substances resulting from the chemical reaction.
    c. The substances on the left side of the chemical equation.
    d. None of the above
5. Not recovering all of the reaction product from the reaction mixture
    a. Results in a falsely low actual yield.
    b. Results in a falsely high actual yield.
    c. Results in a falsely low theoretical yield.
    d. Results in a falsely high theoretical yield.
6. A limiting reactant is named as such because
    a. It is present in greater amounts then the non-limiting reactant.
    b. It creates less product then the non-limiting reactant.
    c. It limits the amount of product that the reaction can produce.
    d. After the reaction has gone to completion, it is still present in excess amounts.

Appendix G

Demographic Questionnaire

Name ________________________

1. Please provide your e-mail address ________________________

2. What is your date of birth? ____________________

3. What year in school are you (circle one).

Freshman Sophomore Junior Senior Graduate Student Other

4. What is your major?

5. Have you taken any chemistry courses at the high school level or beyond? YES / NO

a. If yes, please fill in the following table

Course Name	**Semester/Date Taken**	**Course Completed?**	**If YES, what grade did you receive?**
		YES / NO	
		YES / NO	
		YES / NO	
		YES / NO	
		YES / NO	

This study involves the use of eye tracking technology in which your gaze patterns will be evaluated while reading chemistry text. Studies have shown that some eye conditions affect the validity of the eye tracking data. The following questions will ask about any eye conditions you may have.

6. Do you were corrective lenses?   YES / NO

a. If yes, will you be wearing glasses or contacts at the time of the study?

b. Do you have any other eye conditions that you feel the researcher should know about? These may include astigmatism, glaucoma, strabismus, etc?

Appendix H

Instructions for Pathfinder Program

Your task in this experiment will involve judging the relatedness of pairs of terms. In making these types of judgments, there are several ways to think about the items being judged. For instance, two concepts might be related because they share common features or because they frequently occur together. While this kind of detailed analysis is possible, my concern is to obtain your initial impression of "overall relatedness." Therefore, please base your ratings on your first impression of relatedness.

Page break

Each pair of concepts will be presented on the screen along with a "relatedness" scale. You are to indicate your judgment of relatedness for each pair by pressing a key on the keyboard. If you feel that the concepts are not related at all press "1" on the keyboard. If you feel the concepts are highly related you would press an "8" or a "9". You can think of these numbers as points along a "relatedness" scale, with higher numbers representing greater relatedness. Upon responding, a bar marker will move directly above the number you pressed. If you wish to change your response, simply press another number. When you are satisfied with the rating you have given, press the SPACE BAR to enter your response. Following this, the next pair of items to be rated will be displayed.

Page break

Several pairs of concepts will be shown. If at any time you feel like taking a break tell the researcher and they will save your progress.

Now the complete list of concepts will be presented. This is done to give you a general idea of the scope of the concepts you will be rating.

<u>Practice Pathfinder Task</u>
bear
cat
dog
fox
lion
tiger
wolf

Appendix I

Sample Rubric for Grading General Summaries

The Formation of Ions

The nucleus of an atom is unchanged by ordinary chemical processes, but atoms can readily gain or lose electrons. Electrons can be removed from or added to a neutral atom. The charged particle is called an ion. An ion with a positive charge is called a cation; a negatively charged ion is called an anion. The sodium atom for example, which has 11 protons and 11 electrons, easily loses one electron. The cation has 11 protons and 10 electrons, so it has a net charge of 1+. The net charge on an ion is represented by a superscript. The superscripts +, 2+, and 3+ indicate a net charge resulting from the loss of one, two, or three electrons, respectively; -, 2-, and 3- result from the gain of one, two, or three electrons respectively.

The Formation of Ions Rubric						
Content Score: The presence of each of the following statements from the passage is worth 2 pts. Score 1 pt. if the content in the statement is only partially present or contains false information. Score 2 pts if the content is fully present and correct. The statement does not need to be written verbatim but must contain the complete concept or idea presented in the original sentences. It is not important that the examples be exactly the same if included. This score depends only on the content not on the correct terminology. Absence of the statement earns a score of 0. The order in which the sentences appear does not matter in this grading. Only the fact that each statement is or is not present should be considered.						
Statement	The nucleus of an atom is unchanged by ordinary chemical processes, but atoms can readily gain or lose electrons.	The charged particle is called an ion.	An ion with a positive charge is called a cation; a negatively charged ion is called an anion.	The sodium atom for example, which has 11 protons and 11 electrons, easily loses one electron resulting in a cation with a net charge of 1+.	The net charge on an ion is represented by a superscript. Positive superscripts indicate a net charge resulting from the loss of electrons; negative superscripts indicate a net charge resulting from the gain of electrons. **(The second part of this statement can be indicated by the students through examples).**	**Total Score:**
Score						

Terminology: The students will receive **one** terminology score for the whole paragraph. Review the student's use of scientific terminology and assign a score based on the following criteria. Again, these criteria apply the paragraph as a whole, not individual statements.							
Criteria	Score 0 pts if the student uses no scientific terminology or left the text box blank.	Score 1 pt if the student used terminology but used all of the terms incorrectly.	Score 2 pts if the student used 2-3 terms either with all used correctly or one term mistaken.	Score 3 pts if the student used 3-5 terms either with all terms correctly or 1-2 terms mistaken.	Score 4 pts if the student used over 5 terms with 1-2 mistakes.	Score 5 pts if the student used over 5 terms correctly with no mistakes.	**Total Score:**
Score							

Appendix J

Instructions for Creation of General Summary

1. You will now be shown the title from a chemistry passage you read yesterday.
2. For instance, the example problem from yesterday was called
   The Tooth
1. In the space provided please type a summary of the chemistry passage that was connected with the given title. This summary should contain the main topic as well as any additional pieces of information you feel were important to the passage.
2. Once you have finished typing your summary press the → key to progress to the next page.
3. You will complete this process for all six chemistry passages.

References

Albrecht, J. E., & O'Brien, E. J. (1991). Effects of centrality on retrieval of text-based concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 932.

Albrecht, J. E., O'Brien, E. J., Mason, R. A., & Myers, J. L. (1995). The role of perspective in the accessibility of goals during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 364.

American Association for the Advancement of Science (AAAS). (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

American Chemical Society Division of Chemical Education Examinations Institute. (1992). *ACS test-item bank for general chemistry.*

Anderson, C. W. (1999). Inscriptions and science learning. *Journal of Research in Science Teaching, 36*, 973.

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

Anderson, J. R. (2005). Human symbol manipulation within an integrated cognitive architecture. *Cognitive Science, 29*, 313.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press.

Anderson, J. R. (2010). *Cognitive psychology and its implications*. New York: Worth Publishers.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Quin, Y. (2004). An integrated theory of mind. *Psychological Review, 111*, 1036.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory.* Washington, D.C.: Winston.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2*, 396.

Anderson, J. R., Taatgen, N. A., & Byrne, M. D. (2005). Learning to achieve perfect timesharing: Architectural implications of Hazeltine, Teague, and Ivry (2002). *Journal of Experimental Psychology: Human Perception and Performance. 31(4),* 749-761.

Anderson, R. C. (1985). Role of the reader's schema in comprehension, learning, and memory. In H. Singer, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (pp. 372). Newark, DE: International Reading Association.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 89). New York: Academic Press.

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556-559.

Bartlett, F. (1932). *Remembering.* Cambridge University Press.

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

Bitner, B. L. (1991). Formal operational reasoning modes: Predictors of critical thinking abilities and grades assigned by teachers in science and mathematics for students in grades nine through twelve. *Journal of Research in Science Teaching, 28*(3), 265-273.

Bond, E. A. (1941). Tenth-grade abilities and achievements. *Teachers College Record, 42*(7), 641-643.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts for memory in text. *Cognitive Psychology, 11*, 177.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 61*, 717.

Britton, B. K., & Gulgoz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology, 83*(3), 329-345.

Brooks, F. P., Jr. (1962). Architectural philosophy. In W. Buchholz (Ed.), *Planning a computer system* (pp. 5). New York: McGraw-Hill.

Brown, T. L., LeMay, H. E., Bursten, B. E., & Burdge, J. R. (2003). *Chemistry: The central science* (9th ed.). Upper Saddle River, NJ: Prentice Hall.

Bunce, D. M., & Hutchinson, K. D. (1993). The use of the GALT (group assessment of logical thinking) as a predictor of academic success in college chemistry. *Journal of Chemical Education, 70*(3), 183-187.

Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. *Psychological Review, 108*, 847.

Cain, K., & Oakhill, J. (2004). Individual differences in the inference of word meanings from context: Influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology, 96*(4), 671-681.

Camara, W. J., & Echternacht, G. (2000). *The SAT I and high school grades: Utility in predicting success in college.* (No. 10). The College Board: Office of Research and Development.

Chi, M. T. H. (2006). Laboratory methods for assessing experts' and novices' knowledge. In *The cambridge handbook of expertise and expert performance* (First ed., pp. 167) Cambridge University Press.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics probelms by experts and novices. *Cognitive Science, 5*(2), 121-152.

Chomsky, N. (2006) In Jovanovich B. (Ed.), *Language and mind*. New York: Harcourt.

Clark, H. H. (1974). Semantics and comprehension. In R. A. Sebeok (Ed.), *Current trends in linguistics* (pp. 1291). The Hague: Mouton.

Cognitive Fun. (2009). *Visual forward digit span.* Retrieved 3/15, 2011, from http://cognitivefun.net/test/7.

College Board. (2011). *The college board.* Retrieved 3/15, 2011, from http://sat.collegeboard.org/register?s_kwcid=TC|7002|the%20college%20board||S|p|7081104524.

Craig, M. T., & Yore, L. D. (1996). Middle school students' awareness of strategies fro resolving reading comprehension difficulties in science reading. *Journal of Research in Development in Education, 29*, 226.

Cromley, J. G. (2009). Reading achievement and science proficiency: International comparisons from the programme on international student assessment. *Reading Psychology, 30*, 89-116.

Cromley, J. G., & Snyder-Hogan, L. E. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 102*(3), 687-700.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrica, 16*, 297.

D'Agostino, R., Stephens, M. (1986). *Goodness-of-Fit-Technique.* New York: Marcel Deker, Inc.

DeBoer, G. E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching, 37*, 582.

deCastell, S., Luke, A., & MacLennan, D. (1986). On defining literacy. In S. deCastell, A. Luke & K. Egan (Eds.), *Literacy, society, and schooling* (pp. 3). Cambridge: Cambridge University Press.

Diekhoff, G. M. (1983). Testing through relationship judgements. *Journal of Educational Psychology, 75*(2), 227-233.

Einstein, G. O., McDaniel, M. A., Owen, P. D., & Cote, N. C. (1990). Encoding and recall of texts: The importance of material appropriate processing. *Journal of Memory and Language, 29*(5), 566-581.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211-245.

Fletcher, C. R., & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and Language, 27*, 235-244.

Ford, C., Yore, L. D., & Anthony, R. J. (1997). *Reforms, visions, and standards: A cross-curricular view from an elementary school perspective*

Frederiksen, C. H. (1975). Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology, 7*, 371.

Gibson, E. J. (1969). *Principles of learning and development*. New York: Meredith.

Glanzer, M., & Ehrenreich, S. L. (1979). Structure and search of the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 18*(4), 381.

Glover, J. A., Bruning, R. H., & Plake, B. S. (1982). Distinctiveness of encoding and recall of text materials. *Journal of Educational Psychology, 74*(4), 522-534.

Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology, 83*(1), 88-96.

Gonzalvo, P., Canas, J. J., & Bajo, M. T. (1994). Structural representation in knowledge acquisition. *Journal of Educational Psychology, 86*(4), 601-616.

Goodman, K. S. (1985). Reading: A psycholinguistic guessing game. In H. Singer, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (pp. 259). Newark, DE: International Reading Association.

Guindon, R., & Kintsch, W. (1984). Priming macropropositions: Evidence for the primacy of macropropositions in memory. *Journal of Verbal Learning and Verbal Behavior, 23*, 508.

Haas, C., & Flower, L. (1988). Rhetorical reading strategies and the recovery of meaning. *College Compositino and Communication, 39*, 30.

Hair, J., Anderson, R., Thatham, R., Black, W. (1998). *Multivariate Data Analysis*. New York: Macmillan.

Hand, B. M., Prain, V., & Yore, L. D. (2001). Sequential writing tasks' influence on science learning. In P. Tynjala, L. Mason & K. Lonka (Eds.), *Writing as a learning tool: Integrating theory and practice* (pp. 105). The Netherlands: Kluwer: Dordrecht.

Hanrahan, M. (1999). Rethinking science literacy: Enhancing communication and participation in school science through affirmational dialogue journal writing. *Journal of Research in Science Teaching, 36*, 699.

Hazeltine, E., Teague, D., & Ivry, R. B. (2002). Simultaneous dual-task performance reveals parallel response selection after practice. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 527.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychopysics, 57*(6), 787-795.

Howes, D. H., & Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology, 41*(6), 401.

Huba, G. and Bentler, P. (1980). The role of peer and adult models for drug taking at different stages in adolescence. *Journal of Youth and Adolescence, 9* (5), 449-465.

Hull, C. L. (1952). *A behavior system: An introduction to behavior theory concerning the individual organism*. New Haven, CT: Yale University Press.

Hurd, P. D. (1998). Scientific literacy: New minds for a changing world. *Science Education, 82*, 407.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329-354.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Old Tappan, NJ: Allyn and Bacon.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist,* 294-303.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge, UK: Cambridge University Press.

Kintsch, W. (2004). The construction-integration model of text comprehension and its implications for instruction. In R. B. Ruddell, & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1270-1328). Newark, DE: International Reading Association.

Kintsch, W., Kozminsky, E., Streby, W. J., McKoon, G., & Keenan, J. M. (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior, 14*, 196-214.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394.

Klein, P. D. (2006). The challenges of scientific literacy: From the viewpoint of second-generation cognitive science. *International Journal of Science Education, 28*(2-3), 143-178.

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT for predicting first-year college grade point average.* No. College Board Research Report No. 2008-5). New York: The College Board.

Korpan, C. A., Bisanz, G. L., Bisanz, J., & Henderson, J. M. (1997). Assessing literacy in science: Evaluation of scientific news briefs. *Science Education, 81*, 515.

Kotz, J. C., Treichel, P. M., & Weaver, G. C. (2006). *Chemistry and chemical reactivity* (6th ed.) Thomson Brooks/Cole.

Kozminsky, E. (1977). Altering comprehension: The effect of biasing titles on text comprehension. *Memory and Cognition, 5*(482), 490.

Long, D. L., Golding, J. M., & Graesser, A. C. (1992). A test of the on-line status of goal-related inferences. *Journal of Memory and Language, 31*(5), 634-647.

Lorch, R. F., & Lorch, E. P. (1996). Effects of organizational signals on free recall of expository text. *Journal of Educational Psychology, 88*(1), 38-48.

Mayer, R. E. (1983). *Thinking, problem solving, cognition*. New York, San Fransisco: W.H. Freeman and Company.

Mayer, V. J. (1997). Global science literacy: An earth system view. *Journal of Research in Science Teaching, 34*, 101.

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

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81.

Miller, J., & Kintsch, W. (1980). Readability and recall of short term prose passages: A theoretical analysis. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 335-354.

Murray, J. D., & Burke, K. A.Activation and encoding of predictive inferences: The role of reading skill. *Discourse Processes, 35*(2), 81-102.

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy of Sciences.

National Research Council. (2007). *Rising above the gathering storm: Energizing and employing america for a brighter future.* Washington DC: Author.

Newell, A. (1990). *Unified theories of cognition.* Cambridge, UK: Harvard University Press.

Norman, D. A., & Rumelhart, D. E. (1975). *Explorations in cognition*. San Fransisco, CA: Freeman.

Norman, O. (1998). Marginalized discourses and scientific literacy. *Journal of Research in Science Teaching, 35*, 365.

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

Norris, S. P. (1992). *The generalizability of critical thinking.* New York: Teachers College Press.

Norris, S. P., & Phillips, L. M. (1987). Explanations of reading comprehension: Schema theory and critical thinking theory. *Teachers College Record, 89*, 281.

Norris, S. P., & Phillips, L. M. (1994). The relevance of a reader's knowledge within a perspectival view of reading. *Journal of Reading Behavior, 26*, 391.

Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education, 87*, 224.

Nurrenbern, S. C. (2001). Piaget's theory of intellectual development revisited. *Journal of Chemical Education, 78*(8), 1107.

Olson, D. R. (1994). *The world on paper*. Cambridge: Cambridge University Press.

O'Reilly, T., & McNamara, D. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high stakes" measures of high school students' science achievement. *American Educational Research Journal, 44*, 161-196.

Otero, J., Leon, J., & Graesser, A. (2002). *The psychology of science text comprehension.* Mahwah, NJ: Erlbaum.

Ozkal, K., Tekkaya,C., Sungur, S., Cakiroglu, J., and Cakiroglu, E. (2011). Elementary Students' Scientific Epistemological Beliefs in Relation to Socio-Economic Status and Gender. *Journal of Science Teacher Education*, *22* (2), 155-127.

Ozuru, Y., Dempsey, K., & McNamara, D. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction, 19*, 228-242.

Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory and Cognition, 29*(2), 344-354.

Pashler, H. (1990). Do response modality effects support multiprocessor models of divided attention? *Journal of Experimental Psychology: Human Perception and Performance, 16*(4), 826-842.

Pashler, H. E. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220.

Pathfinder 6.3 (2011). *Pathfinder 6.3 software.*

Pellettieri, A. J. (1955). Time-saving reading strategies in chemistry. *Journal of Chemical Education, 32*(11), 591.

Phillips, L. (2002). Making new and making do: Epistemological, normative and pragmatic bases of literacy. In D. R. Olson, D. K. Kamawar & J. Brockmeier (Eds.), *Literacy and conceptions of language and mind.* (pp. 283-300). Cambridge, UK: Cambridge University Press.

Piaget, J. (1972). *Psychology and epistemology: Towards a theory of knowledge.* Penguin.

Pichert, J., & Anderson, R. C. (1977). Taking different perspectives on a story. *Journal of Educational Psychology, 69*, 309.

Pressley, M., & Wharton-McDonald, R. (1997). Skilled comprehension and its development through instructions. *School Psychology Review, 26*, 448-467.

Pylyshyn, Z. W. (1984). *Computation and cognition*. Cambridge, UK: Mit Press.

Quillian, M. R. (1966). *Semantic memory.* Cambridge, MA: Bolt, Beranak, and Newman.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372 -422.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E-Z reader: A cognitive-control, serial attention model of eye-movement behavior during reading. *Cognitive Systems Research, 7*, 4 -22.

Roadrangka, V., Yeany, R. H., & Padilla, M. J. (1983). The construction and validation of group assessment of logical thinking (GALT). Dallas, TX.

Ryan, J.J.; & Ward, L.C. (1999). Validity, reliability, and standard errors of measurement for two seven-subset short forms of the Wechsler Adult Intelligence Scale-III. *Psychological Assessment, 11(2),* 207-211.

Schallert, D. L. (1976). Improving memory for prose: The relationship between depth of processing and context. *Journal of Verbal Learning and Verbal Behavior, 15*, 621.

Schmuckler, J. S. (1983). Books in communicating chemical knowledge. *Journal of Chemical Education, 60*(7), 611.

Schumacher, E. H., Seymore, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological Science, 12*, 101.

Schvaneveldt, R. W. (1990). *Pathfinder associative networks*. Norwood, NJ: Ablex Publishing Corporation.

Schvaneveldt, R. W., & Durso, F. T. (1981). General semantic networks. *Paper Presented at the Annual Meeting of the Psychonomic Society,* Philladelphia, PA.

Schvaneveldt, R. W., Durso, F. T., Goldsmith, T. E., Bree, T. B., Cooke, N. M., & De Maio, J. C. (1985). Measuring the structure of expertise. *International Journal of Man-Machine Studies, 23*, 699 -728.

Shen, B. S. P. (1975). Science literacy. *American Scientist, 63*, 265.

Shortland, M. (1988). Advocating science: Literacy and public understanding. *Impact of Science on Society, 38*, 305.

Silberberg, M. S. (2000). *Chemistry: The molecular nature of matter and change* (2nd ed.) McGraw Hill.

Singer, M. (1994). Discourse inference processes. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics.* (pp. 479 -515). San Diego, CA: Academic Press.

Spearritt, D. (1972). Identification of subskills of reading comprehension by maximum likelihood factor analysis. *Reading Research Quarterly, 8*(1), 92 -111.

Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction time experiments. *American Scientist, 57*, 421- 457.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). New York: Routledge.

Tabachnick, B. G., & Fidell, L. S. (1996). Using multivariate statistics (3rd ed.). New York: HarperCollins.

The College Board. (2005). *SAT*

Thibadeau, R., Just, M. A., & Carpenter, P. A. (1982). A model of the time course and content of reading. *Cognitive Science, 6*, 157-203.

Tobii Technology. (2010). *Tobii studio 2.X: Software release 2.2.*Tobii Technology AB.

Turmo, A. (2004). Scientific Literacy and Socio-economic Background among 15-year-olds—A Nordic Perspective. *Scandinavian Journal of Educational Research. 48* (3), 287-305.

van Dijk, T. A. (1980). *Macrostructures*. Mahwah, N.J.: L. Erlbaum, Associates.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Wellington, J., & Osborne, J. (2001). *Language and literacy in science education*. Buckingham, UK: Open University Press.

Wilson, J. T., & Neubauer, C. (1988). Reading strategies for improved student work in the chem lab. *Journal of Chemical Education, 65*(11), 996.

Wilson, K. (1983). *A review of research on the prediction of academic performance after the freshman year.* No. College Board Research Report No. 83-2). New York: College Entrance Examination Board.

Yore, L. D., Craig, M. T., & Maquire, T. O. (1998). Index of science reading awareness: An interactive constructive model, text verification, and grades 4-8 results. *Journal of Research in Science Teaching, 35*, 27.

Yore, L. D., & Treagust, D. F. (2006). Current realities and future possibilities: Language and science literacy - empowering research and informing instruction. *International Journal of Science Education, 28*(2-3), 291.