EFFECTIVENESS OF NOTETAKING, SELF-QUESTIONING AND SUMMARIZING STRATEGIES ON LEARNING FROM DIAGRAMS

A Thesis in
Educational Psychology

by

Ya-Ling Lan

© 2008 Ya-Ling Lan

Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

August 2008
The thesis of Ya-Ling Lan was reviewed and approved* by the following:

Rayne A. Sperling  
Associate Professor of Education  
Professor-in-Charge: Educational Psychology  
Thesis Advisor

Jonna Kulikowich  
Professor of Education

Kathy Ruhl  
Professor of Education  
Head of the Department of Educational and School Psychology and Special Education

*Signatures are on file in the Graduate School
ABSTRACT

The primary purpose of this experimental study was to evaluate the hypothesis that comprehension strategies known to facilitate expository text comprehension were also effective in assisting students’ learning from complex diagrams. The study was conducted in an online environment. The conditions we investigated in this study were a control condition (C) and three experimental conditions: notetaking (N), self-questioning (Q), and summarizing (S) strategies. Two hundred and twenty-eight undergraduate students logged onto a website and were randomly assigned to one of the four conditions to study a text and a normal distribution diagram and to employ the assigned strategies. Upon completion of reading the diagram with the text, all participants took two learning outcome measures: a multiple-choice recognition test and a free recall task. The amount of time that participants spent on the learning materials and testing tasks was also recorded. Overall, results indicated that no significant effect was found among conditions on the multiple-choice recognition test, but a significant difference was found among conditions on the free recall task. Students in the notetaking condition (N) significantly recalled more idea units than students in the control condition (C). The current findings were consistent with previous research that notetaking facilitated students’ knowledge as measured by free recall. Results also indicated no time difference among conditions. Data revealed significant positive correlations between time and learning outcome variables. Our results point to important conclusions, implications, and directions for future research.
# TABLE OF CONTENTS

LIST OF FIGURES ................................................................................................................. vi

LIST OF TABLES ................................................................................................................... vii

ACKNOWLEDGEMENTS ..................................................................................................... viii

Chapter 1 Introduction ............................................................................................................. 1

   The Importance of Diagrams for Learning from Text ......................................................... 2
   What Makes Diagrams Difficult to Understand? ................................................................. 4
   The Current Study ................................................................................................................ 12
       Notetaking Strategy .......................................................................................................... 12
       Self-Questioning Strategy .............................................................................................. 16
       Summarization Strategy ................................................................................................. 23
   Research Questions ............................................................................................................. 30

Chapter 2 Methodology .......................................................................................................... 34

   Design .................................................................................................................................. 34
   Participants ........................................................................................................................... 35
   Materials and Measures ....................................................................................................... 37
   Procedures ............................................................................................................................ 41
       Control condition ............................................................................................................ 41
       Notetaking Condition .................................................................................................... 42
       Self-Questioning Condition .......................................................................................... 43
       Summarizing Condition ................................................................................................. 44
   Scoring Idea Units in Free Recall Task ............................................................................ 45

Chapter 3 Results .................................................................................................................... 47

   Multiple-Choice Recognition Test ..................................................................................... 47
   Knowledge-Level Items ..................................................................................................... 48
   Application-Level Items .................................................................................................... 49
   ASE-Level Items ................................................................................................................ 49
   Free Recall ......................................................................................................................... 50
       Time ................................................................................................................................ 53

Chapter 4 Discussion, Implications, and Future Research ...................................................... 55

   Discussion ............................................................................................................................ 55
       Multiple-Choice Recognition Items .............................................................................. 55
       Free Recall Measure .................................................................................................... 58
       Time ................................................................................................................................. 58
   Implications ......................................................................................................................... 59
   Future Research .................................................................................................................. 61
Bibliography ............................................................................................................................65

Appendix A  Published Experimental Studies of Notetaking Strategy (1972-2007).............80
Appendix B  Published Experimental Studies of Self-Questioning Strategy (1982-2002) .....84
Appendix C  Published Experimental Studies of Summarization Strategy (1978-2005) ......89
Appendix D  Recruitment Statement ..................................................................................94
Appendix E  Demographic Questionnaire ...........................................................................95
Appendix F  The Diagram of Standardized Test Scores .......................................................96
Appendix G  The Text of Standardized Test Scores ..............................................................97
Appendix H  Free Recall .....................................................................................................99
Appendix I  Multiple-Choice Recognition Test .................................................................100
Appendix J  Activity: Notetaking .......................................................................................104
Appendix K  Activity: Self-Questioning .............................................................................105
Appendix L  Activity: Summarizing ....................................................................................106
Appendix M  Idea Units of Free Recall ............................................................................107
LIST OF FIGURES

Figure 2-1: Study Procedures and the Control and Experimental Conditions .................. 35
Figure 2-2: Screenshot of Demographic Questionnaire ................................................. 37
Figure 2-3: Screenshot of Multiple-Choice Recognition Test ........................................ 40
Figure 2-4: Screenshot of the Diagram and the Text ...................................................... 41
Figure 2-5: Screenshot of Notetaking Condition ......................................................... 43
Figure 2-6: Screenshot of Self-Questioning Condition .................................................. 44
Figure 2-7: Screenshot of Summarizing Condition ...................................................... 45
LIST OF TABLES

Table 1-1: Taxonomy of Graphics.................................................................4
Table 2-1: Participants Demographic Data According to Condition..................36
Table 3-1: Comparative Means on Recognition Test Scores by Condition ............48
Table 3-2: Comparative Means on Free recall by Condition............................50
Table 3-3: Comparative Means on Free recall by Condition Relative to Statistics Courses...52
Table 3-4: Descriptive Data and Pearson-r Correlations Among Outcome Variables. ....54
ACKNOWLEDGEMENTS

First, I would like to express sincere appreciation to Dr. Rayne Sperling, my thesis advisor, professor-in-charge and mentor, for her guidance and support throughout the long period of my study. I cannot express enough gratitude for her encouragement to me in learning computer programming and her guidance in preparation of this manuscript. Also, I am truly grateful to my second reader, Dr. Jonna Kulikowich for her reading and helpful suggestions. Her efforts should not be understated.

Second, I would like to thank the faculty of the Educational Psychology Program for their support and guidance over the past two years. I am extremely fortunate to have worked with such knowledgeable faculty in this program. In particular, I would like to thank Dr. Hoi K. Suen for always being considerate and offering valuable knowledge and experiences to me.

Last, I would like to thank my family for their support throughout my graduate studies. To my husband, Chi-Tsong, thank you for assistance in fulfillment of my dream. I owe much to you. To my parents, thank you for forgiveness when I am not able to see you in the hometown. I thank all the others including my dear friends, Yu-Chu, Won, Mari, Andria, Crystal, Carla, and who have helped me get where I am today.
Chapter 1

Introduction

Textbooks often serve as a primary source for students’ learning. Textbooks provide not only textual information but also visual displays. Visual displays such as diagrams and graphs are not only found in texts but are frequently used for instructional or learning purposes. In science or mathematics, experts and instructors are often so dependent on graphs that they may be unable to demonstrate their work without them. In order to comprehend complex concepts in expository texts, understanding diagrams is often as important as comprehending the written text. Despite the importance of learning from visual displays, unfortunately, students struggle to comprehend complex diagrams (Hegarty & Just, 1989; Holliday, 1973; Vernon, 1950, 1951; Winn, 1993). For example, previous research has established that students find it difficult comprehend displays that present information such as the general structure of human heart (Butcher, 2006; Dwyer, 1975), physical systems (Hegarty & Just, 1989, 1993), weather maps (Lowe, 1993), pump systems (Mayer & Gallini, 1990), geographic maps (Kulhavy, Stock & Kealy, 1993a), historical timelines (Wiley & Ash, 2005), and statistical line charts or bar charts (Guthrie, Weber & Kimmerly, 1993). Learners apparently pay insufficient or inadequate attention to diagrams (e.g., Peeck, 1994), or they do attend to diagrams, but lack comprehension skills to adequately understand them. As a result, learners construct incomplete knowledge or acquire an inadequate representation of complex concepts from diagrams (e.g., Antonietti, 1991; Otero, León & Graesser, 2002). This current study investigated the effectiveness of several comprehension strategies that were hypothesized to support readers’ knowledge acquisition and comprehension of complex diagrams.
The Importance of Diagrams for Learning from Text

One can distinguish two basic forms of representations in the numerous kinds of learning materials students encounter: descriptions and depictions (Schnotz, 2005). Descriptions include texts, mathematical expressions, formulas, and symbols. Texts are the most common kind of description in textbooks. Pictures such as photographs, drawings, diagrams, graphs, and maps are depictive representations. Descriptive representations and depictive representations have different uses for different purposes (Schnotz, 2005). Descriptive representations are more powerful in expressing abstract and broad information. Descriptive representations have the advantage of being informationally complete. On the other hand, depictive representations are more useful to illustrate quantity, shape, size, and orientation in space. In other words, depictive representations may not be informationally complete but they are more useful for drawing inferences, because information can be read off directly of the representations (Kosslyn, 1994).

Experimental evidence suggests that diagrams not only support memory for information (e.g., Mayer, 1989; Mayer & Gallini, 1990) but also promote mental model construction during learning (Butcher, 2006; Schnotz, 2005). When a reader understands a diagram, he constructs multiple mental representations. Two kinds of representations for diagrams are often discussed: visual image and mental model. When a learner reads a diagram, he or she creates a “visual image” as a perceptual representation of the diagram. The visual image represents a perceptual level and is not yet indicative of complete understanding. Based on this visual image, the reader then constructs a mental model of the diagram content. A mental model is a spatial configuration. A mental model is assumed to have an inherent structure that corresponds to the structure of the subject matter with its visual image but excludes the irrelevant details of the visual image. However, a mental model reflects a more flexible and deeper representation that contains additional information from prior knowledge that is not included in the visual image (Schnotz,
Evidence from previous research demonstrated that the addition of diagrams and illustrations to learning materials can support deep understanding because diagrams promote more accurate mental models of the domain knowledge (e.g., Mayer & Anderson, 1992; Mayer & Gallini, 1990). Research evidence also demonstrated that after long retention intervals, retrieving knowledge from a mental model is easier than retrieving knowledge from a propositional representation (Graesser, Millis, & Zwaan, 1997). More recently, Butcher (2006) investigated learning outcomes when students learned about the heart and circulatory system using text only, text with simplified diagrams, or text with more detailed diagrams. The students in the control condition read the text only, while the students in the simplified diagrams condition read the text with a simplified diagram designed to highlight important structural relations; the students in the detailed diagram group read the text with a detailed diagram reflecting a more accurate representation. The findings revealed that students who used diagrams, whether a simplified diagram or a detailed diagram, with text were best able to improve their mental model, whereas students in the text-only condition improved their mental model least.

Another reason that learning from diagrams is important is that diagrams also may encourage inferences generated by participants. Butcher (2006) indicated that participants using diagrams generated more integration inferences than participants using text-only materials. Further, results demonstrated that diagrams supported learners in generating correct inferences. Butcher (2006) suggested that diagrams may have had additional support for integrating information because the diagrams essentially provide a summary of previously learned information whenever a diagram is present during learning. As a visual summary, diagrams may promote students to integrate information or may provide important cues for remembering or understanding (Schnotz, 2005). Robinson and Kiewra (1995) also suggested that using graphic organizers guide readers to construct relations. In sum, diagrams are important for learning from texts. This is because diagrams not only support memory and mental model construction, but also
powerfully serve as visual summaries to depict spatial relations, to integrate information, and to encourage readers’ inferences which are not explicitly expressed in texts.

What Makes Diagrams Difficult to Understand?

There are several basic types of graphics. Lohse, Walker, Biolsi, and Rueter (1991) developed a classification system using empirical data on how users classified graphics (Vekiri, 2002). The current study addresses four common types of graphical displays based on the work of Lohse et al., (1991): diagrams, maps, graphs, and charts. Table 1-1 explains the similarities and differences among basic types of graphics.

Table 1-1: Taxonomy of Graphics.

<table>
<thead>
<tr>
<th>Types and examples</th>
<th>Referents</th>
<th>Characteristics</th>
<th>Level of abstraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagrams</td>
<td>Iconic (e.g., a diagram showing blood flow and the structure of human heart)</td>
<td>Nonarbitrary symbol system because parts of the diagrams correspond to the objects or entities they represent. In iconic diagrams the relative distances of their parts correspond to the relative distances of their referents</td>
<td>Less abstract</td>
</tr>
<tr>
<td></td>
<td>schematic (e.g., a diagram illustrating the water cycle)</td>
<td>Parts, structure, and operation of real objects or abstract entities; processes</td>
<td></td>
</tr>
<tr>
<td>Maps</td>
<td>Geographic maps, route maps (e.g., a subway map), statistical or thematic maps (e.g., weather maps)</td>
<td>Features (or data) and their location (or distribution) in real territory</td>
<td>Less abstract</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nonarbitrary symbol system because the location of map elements correspond to their location in the territory</td>
<td></td>
</tr>
<tr>
<td>Graphs</td>
<td>Line graphs, bar graphs, pie charts</td>
<td>Quantitative data in a way that enables viewers to compare and observe relations among variables</td>
<td>More abstract</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arbitrary symbol system; neither the parts of the display nor their location correspond to the parts and location of their referents</td>
<td></td>
</tr>
<tr>
<td>Charts (network)</td>
<td>Tree diagrams, web-based concept maps, matrices, graphic organizers</td>
<td>Relations among concepts; sequence of events</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arbitrary symbol system; neither the parts of the display nor their location correspond to the parts and location of their referents</td>
<td></td>
</tr>
</tbody>
</table>

Given the classification system of Lohse and colleagues (1991), various categories of graphics differ in arbitrariness and abstraction. For example, diagrams and maps are nonarbitrary symbol systems while graphs and charts are arbitrary symbol systems, because parts of the diagrams correspond to the objects they represent and the location of map elements correspond to their location in the territory. In contrast, graphs and charts are arbitrary symbol systems because neither the parts of graphs and charts nor their location correspond to the parts and location of their referents. Further, within the category of diagrams, we can make distinctions between iconic diagrams and schematic diagrams (Butcher, 2006). Iconic diagrams are signs that are associated with their referent by similarity or by other structural commonality. For example, a drawing of a bird has some similarities with the corresponding referent such as the shape and structures. Thus, iconic diagrams are less abstract and usually depict a close correspondence between the diagram and the concrete object (Hegarty, Carpenter & Just, 1991). Compared to iconic diagrams, schematic diagrams often depict abstract relationships (as in the case of Venn diagrams or flowcharts) and do not preserve the physical relationships present in the source information (Butcher, 2006). For example, Petre and Green (1993) studied the use of electronic circuit schematic diagrams that make heavy use of abstract conventions such as symmetry and proximity for functional association—by novices and experts. Petre and Green’s results overwhelmingly indicated that, unlike experts, novices were unable to make use of abstract notation such as logical groupings and also were unable to determine what was important versus irrelevant in the schematic representations. Hegarty, Carpenter, and Just (1991) also proposed that diagrams become more difficult to interpret as they become increasingly schematic. Schematic diagrams can be difficult to interpret because the reader or learner must be able to understand and make use of abstract visual conventions.

In addition to schematic diagrams, graphs also have a more abstract structural commonality with their referents (Schnotz, 2005). For example, in bar graphs, the heights of the
bars correspond to the values of variables, and usually the sequence of bars corresponds to the sequence of time, or to the sequences of various groups. Generally, iconic diagrams and maps are less abstract and less arbitrary to learners because there is one-to-one correspondence between their elements and the concrete objects. Schematic diagrams, graphs, and charts are more abstract and more arbitrary because they address quantitative data and relations among concepts and do not preserve the physical relationships between their elements and the concrete objects.

However, even though iconic diagrams are less abstract, learners’ comprehension does not always benefit from iconic diagrams (e.g. Butcher, 2006; Parkhurst & Dwyer, 1983). On the one hand, the elements in iconic diagrams are more “real” and hence reduce abstraction for learners. On the other hand, the structural relations in iconic diagrams are not necessarily transparent to all learners. The correspondence between an iconic diagram and the physical object means that increasing complexity in a physical object requires increasing complexity in an iconic diagram. For example, a detailed diagram of the human heart system emphasizes actual physical details such as its shape, orientation, and structure of a real human heart, resulting in an increase of complexity and visual search difficulty (Butcher, 2006). Learners are required to recognize more details in a complex diagram than in a simplified diagram. As such, the use of realism is not always a benefit in iconic diagrams (Parkhurst & Dwyer, 1983). If we simplify this complexity and just emphasize the key function of a human heart in diagram, it may cause another problem. As we remove some of the correspondence between the diagram and the physical object, the diagram increases in additional abstraction. In short, reducing complexity means increasing abstraction, and conversely, decreasing abstraction infers increasing complexity. How to optimize the complexity and abstraction in iconic diagrams to help students benefit most from them is always a dilemma.

In addition to the fact that both abstraction and complexity make diagrams difficult to understand, another angle with which to consider this issue is from cognitive load perspective.
Sweller and his colleagues did several experiments based on cognitive load theory (Sweller, van Merrienboer, & Paas, 1998). Cognitive load theory proposes that learning difficulty may result from the design of instruction and not from the nature of the material to be learned. Some instructional procedures may impose a heavy extraneous cognitive load that interferes with learning. For example, in their studies, students used instructional materials that involved text and displays, such as technical diagrams (Sweller & Chandler, 1994), cross sections of geographic maps (Purnell, Solman, & Sweller, 1991), and geometry diagrams (Mousavi, Low & Sweller, 1995). They also used a variety of measures to assess students’ performance, such as time on task, problem-solving performance; (e.g., errors or number of correct steps used in solution), and memory of facts. Results indicated that nonintegrated materials placed an extraneous cognitive load on students’ learning of new content and their problem-solving performance, because the materials required students to split their attention among the different sources of information (text and diagrams). The overall finding of their studies showed that integrated materials were more effective than nonintegrated materials.

As mentioned above, one of the reasons why diagrams are difficult to understand is related to complexity and abstraction of diagrams. The second reason is because nonintegrated materials cause an increase in cognitive load. In addition to these difficulties, there are other factors that influence student learning from diagrams. In particular, learner characteristics, such as prior knowledge, may mediate learning from diagrams. Hegarty and her colleagues collected data on readers’ eye fixations and found that individual differences in prior knowledge affected comprehension and the quality of readers’ understanding (Hegarty & Just, 1989, 1993). Generally, low-knowledge readers had difficulty learning from diagrams in three ways. First, when low-knowledge readers read a diagram of mechanical devices such as a pulley and gear system, they did not know what parts of the system were relevant to its function, and failed to extract information more selectively (Hegarty & Just, 1989). Second, low-knowledge readers could not
develop a representation of the system from the diagram when the text did not provide all the relevant information (Hegarty & Just, 1989, 1993). Rather, they needed direction from the text to locate and encode information from the diagram (Hegarty & Just, 1989). Third, low-knowledge readers had more difficulty in comprehending parts of the system and integrating information from the text and the diagram (Hegarty & Just, 1993).

In addition to readers’ prior knowledge, visuospatial ability is another learner characteristic known to mediate the effects on learning from diagrams. Visuospatial ability is the ability to mentally generate and transform images of objects and to reason using these imagery transformations (Carroll, 1993). Although research suggests that the differences between high and low visuospatial ability readers may relate to differences in processing strategies, discrete underlying processes (e.g., working memory capacity), or expectations, understanding of the role of visuospatial ability is limited. Mayer and Sims (1994) speculated that diagrams require low-spatial ability students to devote more cognitive resources for the construction of a visual representation in working memory, which reduces the resources they can allocate for building connections between verbal and visual information. It appears that diagrams may be more demanding to process, and thus less beneficial, when students do not have high visuospatial ability.

Another difficulty that individuals encounter is that they do not have dexterous skills for searching information from complex diagrams. In Guthrie and colleagues’ research, students were asked to perform two kinds of tasks: local search and global search. Local search refers to search for facts whereas global search refers to search for information that requires inferences in bar graphs, diagrams, and text. Results indicated that students performed better on local search than on global search tasks (Guthrie, Weber & Kimmerly, 1993).

The cognitive load theory (Sweller et al., 1998) provides a framework for considerations of cognitive economy from multiple representations. Because verbal and visual representations
both constrain and elaborate the interpretations of other representations, understanding of each representation also creates cognitive cost. When more and more representations about one topic are processed, it is possible that additional benefit for comprehension is not worth the additional cognitive costs. This could explain why individuals frequently ignore information sources (Ainsworth, 1999; Sweller et al., 1998).

In summary, when studying from diagrams, students face numerous difficulties. The difficulties include diagrams that are too abstract or too complex; the lack of integration between texts and diagrams causes an increase in cognitive load; or students have insufficient content knowledge, visuospatial ability, or search skills to effectively comprehend diagrams. All of these are challenges to students. As a result, students ignore, fail to distract information, or generate incorrect inferences from diagrams and hence hinder their comprehension. Early studies regarding learning from diagrams demonstrated memory advantages (Dwyer, 1975) but comprehension disadvantage. In contrast, the more recent body of research on multimedia comprehension developed by Mayer and his colleagues (Mayer, 2001), reports a general (although not complete) advantage for memory, and a consistent advantage for deep comprehension when students receive a multimedia (verbal and visual) presentation as opposed to a text-only presentation. In Mayer’s research (e.g., Mayer, 2001, 2003), memory effects were indicated by performance on recall or recognition questions, and comprehension performance was measured by transfer questions that drew from the learning materials. A variety of multimedia learning studies have demonstrated the importance of carefully designed descriptive and depictive presentations and have generated a collection of important principles that inform the effective combination of text and diagram conditions for learning (Mayer, 2001, 2003, 2005).

In addition to carefully-designed descriptive and depictive presentations for students, we also suspect there are strategies available that may help mitigate students’ difficulties in learning from visual displays and accompanying texts. Perhaps the most direct strategy is to simply ask
students to pay attention to diagrams (Hayes & Readence, 1983; Rasco, Tennyson & Boutwell, 1975). However, little previous research indicated a significant increase in picture effectiveness obtained by telling readers to attend to the illustrations (Peeck, 1994). In addition to telling students to attend to diagrams, there may be other available strategies that can be taught to improve students’ comprehension of diagrams. To date, research has not examined the benefits of known reading comprehension strategies for text as applied to the comprehension of diagrams or displays.

Among numerous comprehension strategies, a common strategy addressed in the current study is notetaking. Notetaking is an effective strategy with sufficient empirical support is often used by college students (Katayama & Robinson, 2000; Robinson, Katayama, Beth, Odom, Hsieh, & Vanderveen, 2006; Robinson & Kiewra, 1995; Van Meter, Yokoi, & Pressley, 1994). Numerous researchers have found benefits for notetaking when learning from lectures or online courses (e.g., Robinson et al., 2006) but notetaking as a strategy has not previously been studied in the context of learning from diagrams.

Along with the notetaking strategy, two other strategies were addressed in the current research. The National Institute of Child Health and Human Development (NICHD) (2000) examined 205 research studies that focused on preventing reading difficulties in children, and eventually identified seven comprehension strategies which had sufficient evidence to support efficacy for the strategy. These seven strategies are question asking, monitoring, summarization, question answering, story mapping, graphic organizers, and cooperative grouping (Shanahan, 2005). Among these effective strategies suggested by NICHD (2000), we examined the benefit of two of these strategies for learning from text and diagrams in the current study: self-questioning and summarization.

The self-questioning strategy requires learners to stop periodically as they read and ask themselves questions (Schunk, 2004). Numerous researchers have found benefits for the self-
questioning strategy when learning from text (e.g., Cheung, 1995; King, 1992; 1994; 1997; Klingner, Vaughn & Schumm, 1998). However, the self-questioning strategy has not previously been studied in the context of learning from diagrams.

The last strategy we examined in the current study was the summarization. Summarization requires that learners put the main idea expressed in the text into their own words (Schunk, 2004). Summarization is a strategy often taught to young children (e.g., Armbruster, Anderson & Ostertag, 1987; Doctorow, Wittrock & Marks, 1978; Klingner, Vaughn & Schumm, 1998; Gajria & Salvia, 1992). Numerous researchers have found summarization is an effective strategy when learning from expository texts (e.g., Cordero-Ponce, 2000; Klingner, Vaughn, & Shay, 1998; Mastropieri, Scruggs, Spencer & Fontana, 2003; Philbrick, 2003). However, little research has examined the effectiveness of summarization for learning from diagrams.

Because most high school and college students have practiced notetaking, self-questioning, and summarizing in school settings (King, 1992; Kobayashi, 2006a; 2006b; Robinson & Kiewra, 1995), we believed that instructional prompts would be effective to support students’ use of the selected strategies. What is not clear is the extent to which these strategies known to facilitate text comprehension can be also effective in assisting students to understand diagrams.

In summary, as we described above, students face a variety of difficulties in comprehending diagrams. Part of the problem they face may be that some types of diagrams are too abstract and complex for students, or text and diagrams lack integration and cause increase cognitive load. Part of the problem may be that students lack content knowledge or visuospatial ability. There are strategies that may facilitate text comprehension, but it is unclear whether these strategies can be also effective in assisting students in understanding diagrams. In the current study, we extended comprehension strategy research by providing learners an expository text and
The Current Study

The current on-line study had four conditions: three experimental and one control. The rationale for each is provided here. Each is described more specifically in the methods section. First, participants in the control condition were instructed to read the diagram as for a class. No strategies were required or instructed for this condition. Participants were free to use strategies if they wanted, including no strategies, as they read the diagram.

The strategies we selected in the experimental conditions for comparison in the current work were selected based upon two primary criteria. First, the strategies selected are those known to be effective in facilitating comprehension from expository text generally. Second, the strategies selected included only those known to college learners. As such, it was our expectation that instructional prompts would be effective to support students’ use of the selected strategies. Given these two criteria we selected notetaking, self-questioning, and summarization as the experimental strategy conditions.

Notetaking Strategy

Notetaking was one of the experimental strategy conditions in the current study. Notetaking with paper and pen is very popular as a learning strategy in school settings (Kobayashi, 2006a; 2006b). Notetaking requires learners to construct meaningful paraphrase of the most important ideas express in text (Schunk, 2004). Beginning with Di Vesta and Gray (1972), researchers have identified two functions or effects of note-taking: encoding and external
storage. The encoding effect results from the note-taking process, while the external storage effect results from the activity of reviewing notes (Di Vesta & Gray, 1972). That is, the encoding function of notetaking suggests that students learn as they take down notes. The external storage function suggests that notetakers benefit from the review of notes they have previously recorded. Considering the practical utility of paper and pen, it seems unlikely that traditional notetaking will be completely replaced by electronic learning technologies in the near future. Rather, mixed use will continue (Haas, 1999). More than thirty years later, researchers still advocate the importance and potential of notetaking and reviewing for school learning (Kobayashi, 2006a). Researchers in several recent studies have examined the effectiveness of computer-based note taking. Katayama and Grooks (2003) found that computer-based partial notes were better than were complete notes on application measures, after a 1-week delay. Katayama, Shambaugh, and Edmonds (2005) compared copy-and-paste versus typed notetaking and found that typed notes were better after one-week delay on transfer tests. Igo, Bruning, and McCrudden (2005) also investigated the potential of copy-and-paste web-based note taking. They found that that type of note taking can be enhanced by (a) using graphic organizers rather than outlines, (b) providing cues (topics and categories) in the graphic organizers, and (c) restricting the amount of text that students are allowed to paste.

In actual academic situations, students reported that they frequently took notes during class and knew that notetaking was important for learning. However, research results indicated the quantity and quality of students’ notes were not optimal for review. For example, Hartley and Davies (1978) reported that 98% of American students and 84% of British students responded “yes” to the questionnaire item, “I take notes to have review material for examination”. In another study by Carrier and Newell (1984), 91.7% of dental hygiene students agreed to the statement, “Taking notes is important because I can review them”. Van Meter and her colleagues (1994) interviewed 252 undergraduates regarding the role of notetaking. The student responses included
that notetaking directed attention during class, facilitated understanding and organization of materials, and helped doing homework. The notes also served as a memory aid, and a study aid after a class or before a test (Van Meter et al., 1994). These data suggest that students feel notetaking helps them to comprehend class materials and helps them study while reviewing notes. However, Hartley & Cameron (1967) indicated that because most students are poor notetakers, they typically record less than half of the critical ideas presented in lectures. Even when students took comprehensive notes, they often did so in a format that was not optimal for review. Most students took notes by using an outline format (Robinson & Kiewra, 1995). The kinds of complete notes that have received the strongest empirical support are spatial, rather than linear forms (Robinson, 1998). Peverly, Ramaswamy, Brown, Sumowski, and Alidoost (2007) also indicated that the quality of notes was the only predictor of test performance. These findings suggest that use of notetaking is frequently reported by learners. However, the notes that students record may lack both quantity and quality.

Given the above concerns, a vast amount of research concludes that students take notes, as compared with those who do not, they generally comprehend learning materials better. There are several reasons. First, notetaking requires that students selectively attend to the information, and attention assists in encoding. Also, when students review notes, as compared with those who do not, they comprehend better because they can spend additional time on the more important content. Both notetaking functions highlight the selective attention nature of note taking; notes simply direct students to pay more attention to important details and less attention to trivial details presented in lectures or textbooks (Robinson et al., 2006). Further, evidence indicates notetaking in some certain forms such as partial graphic organizers permitted students to learn the most concept relations and apply knowledge to novel situations (Katayama & Robinson, 2000). Thus, notetaking does not just assist students by directing their attention to important information,
but, rather, notetaking helps students notice and apply important across-concept relations. In short, notetaking involves attention-directing and relation-revealing advantages.

Appendix A presents a summary of notetaking research for learning materials, intervention conditions, measures, learners’ academic levels, and main findings. The overall findings illustrate several trends. First, research suggested that students’ notetaking substantially enhanced learning even without any special intervention (e.g., Trasborg, 2005). Second, when comparing outcomes with interventions and without interventions (e.g., Trasborg, 2005; Faber, Morris, & Lieberman, 2000), Kobayashi (2006b) meta-analysis study revealed that overall intervention effect was modest but significantly greater than zero. Third, when comparing intervention effect among various types of instruction, results indicated that providing a framework or instructor’s notes is more effective in the enhancement of notetaking effect than pre-training or verbal instruction only (Katayama & Robinson, 2000; Kobayashi, 2006b). Moreover, partial complete notes in spatial forms were better than notes in linear forms. Further, with respect to learners’ characteristics, students at lower academic levels gained greater benefits from intervention compared to students at a higher academic level (Kobayashi, 2006b). Another variable moderating learning outcomes of notetaking is whether notetakers were provided time for later review (encoding plus external storage) or were not allowed reviewed (encoding only). The overall findings suggested that later review of notes substantially heightened the value of notetaking (Kiewra, 1989; Kiewra, Dubois, Christian, McShane, Meyerhoffer, & Roskelley, 1991).

Notetaking is an effective comprehension strategy for a variety of dependent learning outcomes. Learning outcomes such as free recall, short answer, completion, multiple-choice tests, and problem-solving were most commonly measured by notetaking research. The existing research base indicated notetaking advantages were found on factual-recall tests and recognition performance (Kiewra, 1989), comprehension and retention of expository text (Trasborg, 2005),
and problem-solving and self-explanation (Trafton & Trickett, 2001), but not found on tests measuring higher-order performance (Kiewra, 1989).

Despite the practical importance of notetaking on learning, little is known regarding the effectiveness of notetaking when students study materials alone instead of listening to an instructor’s lectures. In the current study, we examined notetaking as an experimental condition when students independently study complex diagrams. According to literature of comprehension of diagrams, when studying, students ignore, fail to distract information, or generate incorrect inferences from diagrams and hence hinder their comprehension of diagrams. Thus, participants in an experimental condition of the current study were directed to generate notes about the diagram in their own words as they studied class materials. We reasonably expected that notetaking would help students pay more attention to diagrams, select important information from diagrams, pay less attention to trivial details presented in diagrams (Robinson et al., 2006), and hence would improve knowledge and comprehension, especially for low-knowledge learners.

**Self-Questioning Strategy**

Self-questioning is another promising text strategy that promotes comprehension by helping students activate prior knowledge, summarize text, and check their understanding of the materials (Gajria, Jitendra, Sood & Sacks, 2007). Self-questioning requires learners to stop periodically as they read and ask themselves questions (Schunk, 2004). The National Reading Panel Report (NICHD, 2000) suggested that only a handful of strategies have empirical support regarding their effectiveness in helping students learn from text. Self-questioning is one of this limited number of strategies (Shanahan, 2005).

At least two phases of questioning may be discerned. *Raising* and *posing* a question (van der Meij, 1994). In the phase of raising a question, the student becomes aware of a knowledge
deficit or experiences a cognitive conflict, and starts searching for an answer. The search occurs in the mind, as a form of inner speech or inner dialogue. If an answer is not found, a question may be posed and involves the questioner and one or more respondents (e.g., teacher, peers). In the current study, self-questioning refers to the raising of questions by individual students, and not involving the social interaction of posing questions (Janssen, 2002).

Self-questioning in reading and learning has been examined from different theoretical perspectives. Wong (1985) and Rosenshine, Meister and Chapman (1996) distinguish three theoretical perspectives of questioning: active processing, metacognitive, and schema theories.

The first theoretical foundation is **active processing theory**. “The overwhelming majority of self-questioning instructional research studies appears to ensue from a theoretical assumption that for students to be active comprehenders and independent thinkers, they must generate questions that shape, focus, and guide their thinking in their reading” (Wong, 1985, p.228). In this perspective, asking more higher-order or think-type questions leads to increased comprehension (Janssen, 2002).

The second theoretical foundation is **metacognitive theory**. Metacognitive theory entails two instructional implications to self-questioning instructional research: “(a) teaching students to be sensitive to important parts of the text by asking questions…..(b) teaching students to monitor their state of reading comprehension by asking questions” (Wong, 1985, p231). The theory highlights the need for developing an awareness of the mental processes and cognitive functioning that allows students to monitor and self-regulate their comprehension through self-monitoring or self-testing questions such as “Is there anything I don’t understand?”

The third theoretical foundation is **schema theory**. “The instructional implication from schema theory for self-questioning instructional research lies in teaching students to activate relevant prior knowledge through appropriate self-questions to aid processing of prose” (Wong, 1985, p232). However, one's reading comprehension may suffer not from lack of prior knowledge
but from lack of activating it (Bransford, Stein, Vye, Franks, Auble, Mezynski & Perfetto, 1982). Those endorsing schema theory assume that students must be able to activate relevant prior knowledge to build “mental scripts” or “mental images” during reading. These questions activate students’ prior domain-specific knowledge and hence influence understanding (Janssen, 2002).

In short, these theories, general in nature, do not make any distinctions between various kinds of texts or purposes. Most often self-questioning is understood as a specific study strategy designed to enhance students’ comprehension and recall of expository texts about specific subject matter (e.g., biology or history) (Janssen, 2002). Asking and answering questions during learning presumably facilitates students' comprehension by inducing such cognitive activities as focusing attention, organizing the new material, and integrating the new information with existing knowledge (Palincsar & Brown, 1984). Furthermore, self-questioning is also considered to be a metacognitive strategy because it helps learners to check how well they comprehend learning materials (Palincsar & Brown, 1984). Specifically, King (1991, 1992) also attributed the success of the self-questioning procedure to the metacognitive (comprehension-monitoring) nature of self-questioning. In sum, self-questioning was used for both comprehension-fostering and comprehension-monitoring.

We already know the functions of self-questioning from several theoretical perspectives. From a practical perspective, why and how self-questioning can help learning has been also explored. For example, Janssen (2002) reviewed 32 studies published between 1992 and 2000 and concluded three common concepts related to the questioning strategy: “transfer of responsibility,” “authenticity of questions” and “engagement in reading.” Transfer of responsibility suggests that asking their own questions help students become independent readers and direct their own reading and learning process. The second common concept is authenticity of questions. This concept focuses on the questioners’ quest to know, exercise curiosity, or express confusion, interest or discovery in reading. Ultimately, self-generated questions may promote
students’ personal *engagement in reading*. When students pose important questions that address what they need to understand, they become intrinsically motivated (Janssen, 2002). In short, three common concepts suggest that by generating or reflecting on questions, students actively and purposefully engage in the clarification of misunderstanding, interpretation of texts, and generation of more reasonable hypotheses.

A large body in literature has examined self-questioning, focusing on different ability level of learners, question types, or training periods. Appendix B presents a summary of self-questioning studies from elementary school through college students.

One focus showed in Appendix B is that while prompted to self-question, what kinds of questions students ask themselves and what kinds of questions most help students. With respect to narratives, Kooy (1992) inferred that the students predominantly generated why-questions; questions about reasons, causes and consequences of events, states, and actions and about author’s motives and intentions. Kooy’s (1992) inference was consistent with the finding of Trabasso and Magliano (1996) that readers predominantly attempt to explain why events, states and actions occur in a story. In Scardamalia and Bereiter’s (1992) study, they classified students’ questions into four rating scales: Complexity of Search; Interest, Fact/Explanation, and Knowledge Advance. They tested these categories with 5th–6th-grade students and adult raters who judged the students’ questions on the Knowledge Advance Scale. A high level of agreement occurred on questions about students’ basic understanding of topics. Rosenshine and colleagues (1996) reviewed intervention studies and found that different prompts yielded different results. For example, signal words (e. g., who, what, where, when, why, how), generic question stems (e.g., How are . . . and . . . alike?) and story grammar categories (different story elements, e.g. setting, main character, character’s goal) appeared to be the most successful in students’ comprehension.
With respect to questioning on expository texts, King (1994) suggested that both lesson-base questioning group (questioning and explaining based on lesson materials) and experience-based questioning group (questioning and explaining based on experiences) outperformed the control group (unguided questioning). She also indicated that experience-based questions were superior to lesson-based questions. Later, King (1997) compared sequenced-inquiry group, un-sequenced-inquiry group, and non-questioning group. She found sequenced-inquiry group outperformed the other groups on comprehension and integration of material, inferencing and knowledge recall. However, King and her colleagues replicated the early experiment found both sequenced-inquiry group and un-sequenced-inquiry group scored better on knowledge tests than non-questioning group, but on the contrary to their early experiments, no significant differences between sequenced and unsequenced questioning were found (King, Staflieri & Adelgais, 1998).

Similar results to King’s work were revealed by Neber’s study. Neber (1999) suggested that students with the causal questioning training reached higher knowledge levels than the students with the factual questioning training. However, Neber (1999) found students in knowledge-generating questioning condition (training in knowledge-generating questioning) outperformed the students in the process-controlling questioning condition (training in process-controlling, metacognitive questioning), particularly in knowledge construction (Janssen, 2002).

As for differences between student-generated and teacher-generated questions, Vittayarungrangsri (1993) reported that the student-questions group outperformed the teacher-questions group on reading comprehension and on number of questions generated by students. In contrast, no significant differences between peer-assisted procedure and teacher-assisted procedure groups were found in Ezell, Hunsicker, and Quinque’s (1997) study. The other example of student-generated questions not being superior to teacher-generated questions was found in El-Koumy’s study. El-Koumy (1996) investigated on higher education learners and found that the teacher-questioning group scored higher than the student-questioning group.
Overall findings from these previous studies revealed that readers predominantly attempt to generate “why” questions or questions about “basic” understanding of topics while reading stories. While reading expository texts, readers benefited more from experience-based questioning, causal questioning, and student-questioning than from factual-questioning and teacher-questioning.

The effectiveness of self-questioning for varying students’ ability levels is inconclusive. One might predict that self-questioning would be more effective with below-average students, who most need strategies, and least successful with the above-average students, who are already engaging in comprehension-fostering activities. In two studies (e.g., Cheung, 1995; Wong & Jones, 1982), below-average students did make greater gains than did other students in the same studies. Other studies (e.g., Dermody, 1988; MacGregor, 1988), however, indicated that average readers made significantly greater gains than good readers or poor readers in comprehension or knowledge tests. In contrast, there are several studies indicating that above-average students made significantly greater gains than did control students. For example, Scruggs and Mastropieri (1985) suggested when gifted-students used elaborative interrogation, a specific self-questioning strategy, they learned more than non-gifted students. In short, the results in these studies do not support that below-average students benefit more from the training of questioning than above-average students. The effectiveness of self-questioning with varying levels of student ability still remains controversial.

Other previous research addressed the comparison of one versus multiple cognitive strategies. One example of multiple strategy instruction was the reciprocal teaching approach. Reciprocal teaching involved instruction in question-asking, summarizing, clarifying, and predicting. Questioning consumed most (estimated 75%) of the instructional time. In Rosenshine and colleagues’ review, they compared reciprocal teaching and regular instruction studies that used the same procedural prompt and obtained similar results (Rosenshine, Meister, & Chapman,
1996). However, contrary to Rosenshine and colleagues’ review, El-Koumy (1996) found that reciprocal teaching group scored higher than student self-questioning group. Furthermore, Gajria and colleagues (2007) suggested larger effect size for instruction in multiple comprehension strategies when compared to a single strategy through reviewing 29 studies. In other words, the question of whether single cognitive strategy instruction or multiple cognitive strategies instruction yield similar results still remains unclear. In this work, we examined an independent self-questioning strategy condition.

In summary, self-questioning has been viewed from different theoretical perspectives. All these theories assume that student readers should be “active” and ask themselves “good” questions about the materials they read. In previous studies (as noted in Appendix B), students in self-questioning groups have outperformed control groups on reading comprehension tests (Klingner, Vaughn & Schumm, 1998; Penkingcarm, 1992; Cheung, 1995), knowledge tests (King, Staflieri & Adelgais, 1998), knowledge recall (King, 1997; Cheung, 1995), free recall (Rich & Shepherd, 1993), integration of material (King, 1997), inference (King, 1997), question tasks (Rich & Shepherd, 1993), and ability to identify referents in text (Kitajima, 1997). Research on self-questioning, and its impact on reading comprehension is extensive and varied. Most of the studies were conducted with elementary or middle school children but few of them were conducted with college students or adult learners. In addition, there are numerous questions related to the self-questioning strategy that remain unanswered. For example, the benefits of self-questioning for varying student ability levels remain controversial. In the current study, we examined self-questioning as a comprehension strategy condition when college students study materials independently. Additionally, the current study examined self-questioning’s impact on learning from complex diagrams. According to the theories, the effectiveness of self-questioning is attributed to both its cognitive and metacognitive functions. Studies also indicated that students can be trained to ask questions during reading and that such training may lead to significant gains.
in reading comprehension and recall. Thus, participants in a self-questioning condition of the
current study were required to generate questions about the diagram in their own words as they
studied class materials. We reasonably anticipated that self-questioning would help students
actively process the diagrams they read. As such, students using this self-questioning strategy
would be more likely to activate their prior knowledge, be engaged in reading, and have chances
to clarify misunderstanding, and hence improve knowledge acquisition and comprehension than
students who engaged in free study.

**Summarization Strategy**

Another experimental strategy condition in the current study was the summarization
strategy. By definition, summarization requires that learners put the main idea expressed in the
text into their own words (Schunk, 2004). Summarization of reading materials has been found to
enhance comprehension of expository text (e.g., Gajria & Salvia, 1992; Hutton, 2002; Klingner,
Vaughn & Schumm, 1998) and recall of passage content (e.g., Cordero-Ponce, 2000; Elosua,
Garcia-Madruga, Juan, Gutierrez, Luque & Garate, 2002; Wittrock & Alesandrini, 1990).

Among numerous strategy instruction studies, the National Reading Panel Report
(NICHD, 2000) examined 205 research studies and eventually identified seven comprehension
strategies which had sufficient evidence supporting their use. The summarizing strategy was one
of the identified strategies.

How can summarization help learning? A theoretical foundation for summarization is
Wittrock’s model of generative processing (Wittrock, 1990). In this view, comprehension is a
generative process rather than discovery process. Wittrock’s generative learning process includes
four major components: generation, motivation, attention, and memory. By definition, generation
is a fundamental cognitive process in comprehension. “Comprehension involves the reader’s
active generation of these two types of semantic and pragmatic relations: (a) among parts of the
text, and (b) between the text and knowledge and experience” (Wittrock, 1990, p. 348). The
active generation of these two types of relations implies a *motivation* or willingness to invest
efforts in reading. The third element in this model, *attention*, directs the generative process to
relevant information, related stored knowledge, and memory of pertinent experience. The fourth
element, *memory*, includes perceptions, metacognition, abstract knowledge, and concrete
experiences. In other words, peoples’ minds or brains are not passive consumer of information.
Instead, they construct their own interpretations of information and draw inferences from them.
People ignore some information and selectively attend to other information. People also retrieve
information from long-term memory and use their information–processing strategies to generate
meaning from incoming information to organize it, to code it, and to store it in long-term memory.
From such a perspective, students’ knowledge and learning strategies are crucially important
because comprehension depends on what they do with the new information (e.g. summarizing)
and what they think about it (e.g. questioning), and how they related it to their knowledge (e.g.
elaboration).

According to Wittrock (1990), there are numerous ways to stimulate generative processes.
For example, teachers can ask students to construct headings, inferences, main ideas, summaries,
questions, analogies, answers, pictures, and alternative and explanations while they are reading or
after they read the materials (Wittrock, 1990). Reading does not always involve these same
constructive processes. When we make reading more like written composition and engage
generative activities, such as generating summaries or questions, it seems that reading
comprehension can be increased. Over the last 30 years, Wittrock and his colleagues have
completed a series of empirical studies to test implications for teaching of generative processes.
Three applied studies of generative reading comprehension are reviewed here.
In one study with six-grade students (Doctorow, Wittrock & Marks, 1978), those in one experimental group were asked to generate a summary sentence for each paragraph they read, and another experimental group was given paragraph headings to use in the summary sentences they were asked to construct. The control group was only asked to read the text but not generate summaries. Results indicated students who generated summaries sizably and statistically increased their retention and comprehension of the text. Furthermore, the group given headings and asked to generate summaries doubled their retention and comprehension. Another experiment on junior high school students compared three groups: control group, free generation group, and written group. The control group wrote the same sentences generated by their counterparts in the free generation experimental group. Results showed that effects for summarizing occur, indicating that students’ comprehension is influenced by the active process of generating representations for meaning in the materials, “not only by effort or by practice, or by writing, or by additional information in the insert sentences, all of which were controlled by this design” (Wittrock, 1990, p.363). In another series of applied studies in basic skills conducted over a period of two and a half years at four Army bases (Wittrock & Kelly, 1984; from Wittrock, 1990). Soldiers with low reading ability were taught basic strategies including imagery, verbal, summarization, and metacognitive strategies. The soldiers in each of the experimental groups showed statistically significantly greater improvement in reading comprehension than did the soldiers in the control groups. Overall, the findings of these studies often parallel the results of other studies on generative reading comprehension. In short, from a generative processing perspective, comprehension is not the process of transforming a stimulus on a page into a product only; the process does not only involve transforming input to output. Rather, reading comprehension involves generative processes that generate signals, strategies, and plans that relate events to one another and to memory to give them meaning and significance for understanding.
In addition to Wittrock and his colleagues’ work, more recently a large body of research has examined the general effectiveness of summarization training on various kinds of learners, learning environments, and learning tasks. The purposes of these studies were varied and findings are mixed. Appendix C presents a summary of summarization and main idea research. The learning materials, intervention conditions, measures, learners’ abilities or academic levels, and main findings are provided.

A vast amount of experimental studies in Appendix C have two consistent conclusions. First, summarization strategies can be taught (e.g., Armbruster, Anderson, & Ostertag, 1987; Gajria & Salvia, 1992; Mastropieri et al., 2003). Second, compared with students who did not receive training, students who received summarizing training can increase their content knowledge (Elosua et al., 2002; Hutton, 2002) and reading comprehension (Elosua et al., 2002; Gajria & Salvia, 1992; Klingner et al., 1998; Mastropieri et al., 2003). However, there are studies yielding inconsistent results in retention or transfer measures and strategy maintenance. For example, results of Wang’s (2001) study with undergraduate students revealed no significant difference between summarizing strategy condition and structure condition on transfer tasks, but significant results were shown by Gajria & Salvia (1992), indicating summarizing was maintained over time, and students generalized its use to different content areas. Meanwhile, Malone & Mastropieri (1992) found students who were trained to generate a summary sentence for each paragraph outperformed students who received traditional instruction on immediate posttest, near transfer, and far transfer measures. Similarly, Jitendra, Hoppes and Xin (2000) found benefits on a post test measure that required selection and production, was maintained six weeks after training, but in contrast, transfer effects were less robust. In short, except for consistence in comprehension improvement and knowledge acquisition, findings of research investigating effects for summarizing are quite mixed and varied.
In addition to single strategy of summarizing, the effectiveness of multiple-comprehension strategies integrated with summarizing has also been examined. The combined use of summarizing and other comprehension strategies included main idea identification (Elosua et al., 2002), questioning (Brown & Palincsar, 1989; Hogewood, 2004; Klingner, Vaughn & Schumm, 1998), comprehension monitoring (Graves & Levin, 1989; Malone & Mastropieri, 1992), self-questioning (Glover, 2002), self-monitoring (Jitendra et al., 2000; Ferguson; 2001; Malone & Mastroperi, 1992), mnemonic (Grave & Levin, 1989), mnemonic imagery (Glover, 2002), clarification (Brown & Palincsar, 1989; Hess, 2005; Klingner et al., 1998). The national Reading Panel Report (NICHD, 2000) also suggested that the teaching of the combined use of multiple strategies has been most effective in improving reading (Shanahan, 2005).

Among numerous multiple-strategy reading instruction programs, three examples were most discussed by researchers: reciprocal teaching (Brown & Palincsar, 1989), Collaborative Strategic Reading (CSR) program (Klingner, Vaughn & Schumm, 1998), and Concept-Oriented Reading Instruction (CORI) (Guthrie, Van Meter, McCann & Wigfield, 1996; Guthrie, Anderson, Alao & Rinehart, 1999; Guthrie & Wigfield, 2000; Guthrie, Wigfield, Barbosa, Perencevich, Taboada, Davis, Scafiddi & Tonks, 2004).

Reciprocal teaching is an example of multiple strategy instruction. Reciprocal teaching refers to a set of learning conditions in which children "first experience a particular set of cognitive activities in the presence of experts, and only gradually come to perform these functions by themselves" (Brown & Palincsar, 1989, p.123). During reciprocal teaching, students read a passage of expository material and learn and practice four major reading comprehension strategies: generating questions, summarizing, attempting to clarify word meanings or confusing text, and predicting what might appear in the next paragraph. According to Rosenshine & Meister’s (1994) review of 16 reciprocal teaching studies, investigators of 10 studies achieved significant gains in comprehension tests, recall idea units, English tests, questioning, or
summarizing levels by teaching from 2 (questioning and summarizing) to 10 cognitive (four major strategies and other) strategies. Therefore, reciprocal teaching has been found to improve students’ reading expository text from a variety of aspects. However, regarding reciprocal teaching, there remain unanswered questions for future research. For example, reciprocal teaching investigators reported different results between using standardized measures and using experimenter-developed measures. In many of studies, the results were significant when experimenter-developed tests were used and nonsignificant when standardized tests were used. Further, investigators were not clear about which procedural prompts were most useful for teaching cognitive strategies and how many and which strategies are most productive (Rosenshine & Meister, 1994).

The second example of multiple strategy instruction relating summarizing was Collaborative Strategic Reading (CSR) program. CSR was designed by Klingner and his colleagues to improve secondary students’ reading comprehension. This program combined two instructional elements: modified reciprocal teaching and cooperative learning or student pairing. In CSR training, students were taught by the researchers to apply reading comprehension strategies ("preview," "click and clunk," "get the gist," and "wrap up"). This program was not only conducted with students with learning disabilities (Klingner, Vaughn & Schumm, 1998) but also in heterogeneous, culturally and linguistically diverse, general education classroom settings (Klingner, Vaughn & Schumm, 1998). Studies indicated that students in the training condition made greater gains in reading comprehension. Among several strategies, students implemented the clarification and main idea (summarization) strategies the most consistently and effectively (Klingner et al., 1998).

Another example of multiple-strategy instruction was the Concept-Oriented Reading Instruction (CORI) program conducted by Guthrie and his colleagues (Guthrie et al., 2004). CORI provided a classroom context where the multiple strategies of activating background
knowledge, questioning, searching for information, summarizing, and organizing graphically were taught to third-grade children. Different from reciprocal teaching of multiple strategies (e.g., Brown & Palincsar, 1989), the instructional framework for CORI was a combination of strategy instruction and motivation support. Class-level analyses showed that CORI surpassed traditional instruction in reading comprehension (Guthrie et al., 1999; Guthrie et al., 2004), strategy learning (Guthrie et al., 1999; Guthrie et al., 2004), and reading motivation (Guthrie et al., 2000; Guthrie et al., 2004). However, these studies were done with third or fifth graders, and how or whether the findings might generalize to other age groups is still unclear.

As mentioned earlier, summarization strategy should be known to college learners. However, the ability to create summaries develops slowly, and even many high school and adult students have difficulty with this skill (Brown & Day, 1983). The National Reading Panel also reported that readers do not identify main ideas, summarize text, or integrate comprehension strategies to effectively construct meaning (NICHD, 2000). Students all felt familiar with summarizing strategy but whether or how to use summarizing became an issue. Researchers suggested that we need further research on this issue on the generation of summary in exposition, narration, and other types of discourse; we need to study how to teach what a summary does in each of the different types of discourse (Wittrock, 1990; Gajria et al., 2007).

In sum, according to generative processing model, one of the constructive processes is generating summaries. Previous researchers have examined the effectiveness of summarization strategies among different populations, such as struggling learners, with different lengths of target texts and text topics, and explored whether the integration of summarization with other strategies yields differences. However, as described in Appendix C, except for Brown & Day (1983), Cordero-Ponce (2000), Wang (2001), Wittrock and Kelly (1984), Wittrock and Alesandrini (1990), little research has addressed older learners or college students. Moreover, little research had examined the effectiveness of summarization with a complex diagram. In this current study,
we examine summarizing as an experimental study condition in which participants were required to identify main ideas from the diagram and generate a summary in their own words. Because of the generative nature of summarization, we hypothesized that reading a complex diagram and summarizing should result in improvement of understanding and memory of the information from the diagram. We expected that summarizing would help students use their own words and experiences to construct novel sentences. Those novel sentences would not only organize information from the diagram but would also make connections among the concepts and subsequently would relate new information to the learners' prior knowledge; resulting in knowledge acquisition and improved comprehension.

**Research Questions**

Students have difficulty comprehending expository information. Expository information in textbooks that students often encounter includes text and other representations such as diagrams. There are strategies that may facilitate students’ comprehension, but it still remains unclear whether they can effectively benefit understanding from an expository text in conjunction with a complex diagram. In this work we examined whether comprehension strategies known to facilitate expository text comprehension were also effective in assisting students’ learning from diagrams. Specifically we compared students’ use of notetaking, self-questioning, and summarizing strategies. We posed six research questions to address our objectives.

1. Are there differences in students’ recognition test total scores among the control, notetaking, self-questioning, and summarizing conditions?
2. Are there differences in students’ scores of knowledge-level recognition test items among the control, notetaking, self-questioning, and summarizing conditions?
3. Are there differences in students’ scores of application-level recognition test items
among the control, notetaking, self-questioning, and summarizing conditions?

4. Are there differences in students’ scores of ASE-level (a combined level of analysis, synthesis and evaluation) recognition test items among the control, notetaking, self-questioning, and summarizing conditions?

5. Are there differences in the number of idea units students generate during a free recall task among the control, notetaking, self-questioning, and summarizing conditions?

6. Are there differences in the time that students spend in reading, and answering free recall, recognition test items among conditions? Further, is the length of time correlated with learning outcomes?

The first research question addressed whether there are differences in students’ total recognition test scores among the control, notetaking, self-questioning, and summarizing conditions. We expected students in the three comprehension strategy conditions to outperform students in the control condition on the total recognition test score. However, for the comparison among three comprehension strategy conditions, we had no firm hypothesis. Students in the three comprehension strategy conditions might perform either equally or one condition might outperform others. The three expected results were control < notetaking, control < self-questioning, and control < summarizing.

The second through the fourth research questions addressed whether there are differences in students’ scores of knowledge-level, application-level, and ASE-level of recognition test items among the control, notetaking, self-questioning, and summarizing conditions. In knowledge-level and application-level item scores, we expected that students in all of the three comprehension strategy conditions would outperform the control condition but we expected no differences among the three strategy conditions. Therefore the pattern of expected results was control < note taking = self-questioning = summarizing. Regarding the ASE items,
previous research (Kiewra, 1991; Van Meter et al., 1994; Wittrock, 1990) suggests that both notetaking and summarizing comprehension strategies facilitate students’ organizational strategies. Moreover, according to previous research, students used more higher-level questions of critical thinking in their discussion after notetaking or summarizing training (Hess, 2005; Jitendra et al., 1998). Therefore, we hypothesized that students who used notetaking or summarizing strategies would outperform students in the control or self-questioning conditions. We expected the pattern of results for the ASE-level scores to be \textbf{control} = \textbf{self-questioning} < \textbf{notetaking} = \textbf{summarizing}.

The fifth research question examined \textbf{whether there are differences in the number of idea units students generate during a free recall task among the control, notetaking, self-questioning, and summarizing conditions.} There is research evidence to suggest that comprehension strategies improve recall (Brown et al., 1996; Cordero-Ponce, 2000; Elosua et al., 2002; Kiewra, 1991; King, 1997; Rich & Shepherd, 1993). Therefore, we expected all of the three comprehension strategy conditions to outperform the control condition in free recall. Furthermore, since participants in notetaking and summarizing conditions engaged in a notetaking or summary activity, which was similar to free recall, we expected students in these notetaking and summarizing conditions to outperform those in the self-questioning condition. The expected pattern of results was \textbf{note taking} = \textbf{summarizing} > \textbf{self-questioning} > \textbf{control}.

The sixth research question addressed \textbf{whether there are differences in the time that students spend in reading and free recall and recognition testing among conditions?} Further, we were also interested in \textbf{whether the length of time students used correlated to the recognition and recall learning outcomes.} This research question served two purposes. First, we used time as a measure to assure that students were engaged with the materials. Second, by investigating time we were able to ascertain that time on materials alone was not the important factor, over strategy condition, in students’ comprehension. We did not expect that the amount of
time on tasks across conditions would differ. However, we hypothesized that the longer time students engaged in reading, free recall, and taking the recognition test, the higher recognition test total scores they had and more idea units they recalled from the text or the diagram.
Chapter 2

Methodology

Design

This study was a posttest-only control group design. To participate, students were directed to a website where, at their convenience, they engaged in the study. Participants were randomly assigned to one of the four conditions, three experimental (notetaking, self-questioning, or summarizing) or one control. The independent variable was represented by these four conditions. The dependent measures included a multiple-choice recognition test containing knowledge-level items, application-level items, and ASE-level (analysis, synthesis, or evaluation-level) items, and a free recall task. This study was conducted in an on-line learning environment. Each implementation step of this study was shown in Figure 2-1.
Participants

Two hundred and twenty-eight undergraduate students participated in the study. Participants included one hundred and eighty females and forty-eight males recruited from a learning and instruction class at Pennsylvania State University. (Appendix D provides the recruitment statement used in the study.) All participants signed an informed consent form approved by the human subjects review. Participants received a small amount, approximately 1% of total score, of extra credit toward their course grade in exchange for participating in the study. All participants were randomly assigned into one of four conditions: notetaking (n = 59), self-questioning (n = 60), summarizing (n = 46), and control (n = 63). In general, the participants
appeared to be an average group of university students. Of the two hundred and twenty-eight students who participated in the study, 3.10% of the participants were in their first year, 83.33% in their second year, 9.21% were in their third year, 3.51% in their fourth year, and .43% were classified as other. The majority of the participants (96.50%) were between the ages of 18 and 21 with the remaining (3.50%) between 22 and 24. Among the participants, 42.11% had never taken courses in statistics, 43.86% had taken 1 course, and 14.04% had taken 2 or more courses in statistics. Participants’ mean SAT math score was 580.46. Other demographic data are presented in Table 2-1.

Table 2-1: Participants Demographic Data According to Condition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Control</th>
<th>Notetaking</th>
<th>Self-questioning</th>
<th>Summarizing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>12</td>
<td>17</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Female</td>
<td>51</td>
<td>42</td>
<td>51</td>
<td>36</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 to 19</td>
<td>38</td>
<td>35</td>
<td>36</td>
<td>34</td>
</tr>
<tr>
<td>20 to 21</td>
<td>22</td>
<td>21</td>
<td>22</td>
<td>12</td>
</tr>
<tr>
<td>22 to 23</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Others</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Standing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First-year</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Sophomore</td>
<td>51</td>
<td>46</td>
<td>50</td>
<td>43</td>
</tr>
<tr>
<td>Junior</td>
<td>6</td>
<td>8</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Senior</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Others</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Courses taken in statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 course</td>
<td>20</td>
<td>26</td>
<td>24</td>
<td>26</td>
</tr>
<tr>
<td>1 course</td>
<td>30</td>
<td>25</td>
<td>28</td>
<td>17</td>
</tr>
<tr>
<td>2+ course</td>
<td>13</td>
<td>8</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Mean SAT-math score</td>
<td>584.52</td>
<td>588.39</td>
<td>577.46</td>
<td>567.98</td>
</tr>
</tbody>
</table>
Materials and Measures

Demographic Questionnaire

All participants initially completed a demographic questionnaire. The demographic questionnaire collected information including age, gender, semester standing, major, and previously taken statistics courses (see Appendix E and Figure 2-2). The collection of demographic data served two purposes. First, these data provided descriptive information about the sample. Second, the data were used to divide participants into three subgroups: those who had not taken any statistics courses, those who taken one course, and those who had taken two or more than two courses.

Figure 2-2: Screenshot of Demographic Questionnaire.
Diagram

A text and a diagram were the materials used in this study. Participants read the same diagram and text across all conditions. The diagram depicted a normal distribution (See Appendix F for the diagram) downloaded from http://en.wikipedia.org/wiki/Standard_score. The diagram compared the various grading methods in a normal distribution. The information within the diagram included standard deviations, cumulative percentages, percentile equivalents, Z-scores, T-scores, standard nine, percentage in stanine.

Text

The target text was an 1121-word, 11-paragraph passage (See Appendix G for the text). The topic of the target text was “Standardized Test Scores”. The text provided definitions and the calculation procedures for several types of standardized test scores and characteristics of standardized testing. Text topics included measures of central tendency, standard deviation, normal distribution, raw scores, percentiles, z-score, t-score, and stanine scores. The text was presented electronically within a single text box and included section headings, similar to texts found in online learning and as in typical textbooks. The same target text was employed across conditions and instructions for reading were included for each of the condition.

Distractor Task

After reading, all participants completed distractor tasks for 6 minutes. There were two parts in the distractor tasks: a card rotation test and a paper folding test (Ekstrom, French & Harman, 1976). The purpose of the tasks was to distract participants and prevent participants from retrieving knowledge from working memory rather than long term memory.
**Free Recall**

The free recall task then required participants to record everything that they could remember from the diagram or text. All participants received the same instructions: “You just studied a passage and a diagram about standardized test scores. Please type down every thing you can remember from both the passage and the diagram. (See Appendix H for the free recall task.) An unlimited time amount was provided for free recall. However, most participants completed within 10 minutes and all completed within 15 minutes.

**Multiple-Choice Recognition Test**

All participants were next required to take a multiple-choice recognition test (see Appendix I and Figure 2-3) which assessed recognition and comprehension both from the text and the diagram. The knowledge or information of the test content appeared both in the diagram and the text. If the content only appeared in the text but not in the diagram, this content was not tested by the recognition test. For example, the content of “mean calculation” only appeared in the text but not in the diagram, and therefore was not assessed on the recognition test. However, a series of “z-scores” were displayed at the bottom of the diagram and were also discussed in the text, therefore, the recognition test items included the content of ”z-scores”. The recognition test items were designed based on Bloom’s Taxonomy (Bloom, Hastings & Madaus, 1971) and classified into three levels: knowledge-level, application-level, and ASE-level (analysis, synthesis, or evaluation). Knowledge-level items required participants to recall information from the diagram and text. Application-level items required participants to apply laws or use a concept in a new situation. ASE-level items represented combined categories of analysis, synthesis, and evaluation. ASE-level items required participants to separate material into component, integrate
information from several sources, or select the most effective solution. The Multiple-choice recognition test included 3 knowledge-level items, 4 application-level items, and 13 ASE-level (analysis, synthesis, or evaluation-level) items. An example of a knowledge-level item was \emph{A score at the mean of the normal distribution represents a stanine of (a) 0; (b) 1; (c) 5; or (d) 9.} An example of an application-level item was \emph{What is the cumulative percentage of z-score = -2? (a) 2.3%; (b) 15.9%; (c) 50%; or (d) 84.1%.} An example of an ASE-level item was \emph{On a recent standardized test, your student, Becky, had a z-score of 2. What is the closest assessment of her performance compared to the group? (a) Her T-score is negative; (b) Her percentile rank is below 85; (c) Her percentile rank is above 85; or (d) Her stanine score is 4.} The internal consistency reliability of the overall multiple choice instrument in the sample was $\alpha = .65$.

![Figure 2-3: Screenshot of Multiple-Choice Recognition Test.](image)

\emph{Time}

As long as a participant completed the demographic questionnaire and started an activity in a randomly assigned condition, the computer server recorded the participant’s start time. Upon completion of activity, study, and dependent measure tests, the computer server recorded the
participant’s stop time. The amount of time that each participant spent was obtained through the subtraction of the start time from the stop time.

**Procedures**

**Control condition**

After completing the demographic questionnaire, in the control condition, instructions directed participants to read the text and diagram as they would read for a class. Within the online materials, the text was presented on the left side of the text page and the diagram was shown simultaneously to the right (see Figure 2-4 and Appendix F and G for complete content). The instructions for the control condition read:

In this activity, you will read a text and a diagram about standardized test scores. As you read, do your best to understand the main ideas. Your participation in this study will help us to learn more about the strategies for reading from expository text and learning from diagrams. Please read the text carefully to prepare for an upcoming assessment. At the end of the reading activity, you will be asked to answer some test-like questions.

![Figure 2-4: Screenshot of the Diagram and the Text](image)
Once participants in this condition completed reading, they were directed to the distactor tasks. Participants completed the distractor tasks and then were tested on the dependent measures. Participants were not allowed to return to the text or the diagram while taking the dependent measures. After completion of the dependent measures, participants saw the final webpage of the study, which thanked them for their participation.

Notetaking Condition

The notetaking condition focused on prompting participants to use a notetaking strategy. After completing the demographic questionnaire, in the notetaking condition, participants were first asked to read the diagram and to take notes about the diagram in the response box provided. For this page, the diagram was presented on the left hand side while the text box was shown on the right hand side of the screen (see Figure 2-5 and Appendix J). The instructions on the top of this on-line page prompted learners:

In this activity, you will study a diagram about standardized test scores. Please take notes about the diagram similar to the notes you take when you study class materials. You may write words or sentences in the response box on the right side of this page. You may refer to the diagram while you are taking notes.

As they read the diagram, participants typed their notes in their own words into the response box. After completion of the diagram reading and notetaking, participants were directed to read the text and diagram. Participants were provided a diagram in conjunction to a text on the on-line page. The text topic was standardized test scores and the diagram was as the same diagram while they took notes. As in the control condition, the diagram was shown on the left side and the text was shown on the right side of the on-line page. Reading instruction in this condition was similar to the instruction in the control condition. After completion of reading the diagram and text, participants were directed to the distractor tasks and the dependent measures. They were not
allowed to return to the diagram and text page or to their responses while taking the tests. The final web page of the study thanked them for their participation.

Figure 2-5: Screenshot of Notetaking Condition.

**Self-Questioning Condition**

The self-questioning condition focused on prompting participants to use a self-questioning strategy. After completing the demographic questionnaire, in the self-questioning condition, participants were first asked to read the diagram and then generate questions about the diagram in the response box on the screen (see Figure 2-6 and Appendix K for complete content). The diagram was presented on the left side while the response box was shown on the right side of the screen. The instruction on the top of this on-line page prompted learners:

In this activity, you will study a diagram about standardized test scores. Please write down questions you would ask yourself to help yourself understand the diagram as if you were studying for a class. You may generate questions in the response box of this page. You may refer to the diagram while you are writing questions.
After completion of reading and self-questioning, participants read the text in conjunction with the diagram, completed the distactor tasks, and finished the dependent measures as did participants in the other conditions. As in the other strategy conditions, students were not allowed to return to the diagram and text page or to their responses while taking the tests. The final web page of the study thanked them for their participation.

Figure 2-6: Screenshot of Self-Questioning Condition.

**Summarizing Condition**

The summarizing condition focused on prompting participants to use a summarization strategy. After completing the demographic questionnaire, in the summarizing condition, participants were asked to read the diagram and then make a summary about the diagram in the response box on the screen. The diagram was presented on the left side while the response box was shown on the right side of the screen (See Figure 2-7 and Appendix L). The instruction on the top of this on-line page prompted learners:

In this activity, you will study a diagram about standardized test scores. Please **summarize what you learned** from the diagram as you might summarize information when you were learning in a class. You may write words or
sentences in the response box of this page. You may refer to the diagram while you are summarizing.

After completion of diagram reading and summary construction, participants read the text in conjunction with the diagram, completed the distactor tasks, and finished the dependent measures as did participants in the other conditions. Participants were not allowed to return to the diagram and text page or to their responses while completing the tests. The final web page of the study thanked them for their participation.

Figure 2-7: Screenshot of Summarizing Condition.

Scoring Idea Units in Free Recall Task

The free recall task was scored by counting the number of idea units written by participants. The scoring rubric consisted of two categories of idea units (See Appendix M). The first category of idea units was defined as idea units could be recalled from the text. For example, unit 06, “One of the most commonly used central tendency measures is the mode”, was one of the idea units recalled from the text because the concept of “mode” was described by the text but not by the diagram. The second category was defined as the idea units that could be recalled either
from the text or from the diagram because these idea units were presented both in the text and the diagram. For example, unit 25, “The characteristics of a normal distribution include a symmetric distribution” could be recalled either from the text or from the diagram. The researcher reviewed all responses to the free recall task and counted the number of idea units recorded by each participant.
Chapter 3

Results

Chi-Square tests were first conducted to assure that the number of statistics courses that participants had taken were equally likely across conditions. Results indicated that, for the first subgroup who never took statistics, the percentages were equally likely for conditions. \(\chi^2(3, n=96) = 1.00, p=.80\). Similarly, for the second subgroup who took one statistics course, the percentages were equally likely for conditions, \(\chi^2(3, n=100) = 3.92, p=.27\); for the third subgroup who took two or more statistics courses, the percentages were equally likely for conditions. \(\chi^2(3, n=32) = 6.25, p=.10\). In short, the Chi-Square test results indicated no significant differences in the number of statistics courses that participants had taken by condition prior to treatment.

Multiple-Choice Recognition Test

The first research question of the current study addressed whether there were differences in students’ total recognition test scores among the control, notetaking, self-questioning, and summarizing conditions. Table 3-1 presents the means and standard deviations for total recognition test score by conditions. The findings of descriptive statistics revealed, for multiple-choice recognition test total scores, as we expected, students in the notetaking and summarizing conditions had higher mean recognition test scores than those in the control condition. Specifically, those participants in the summarizing condition scored highest among conditions. However, unexpectedly, students in the self-questioning condition did not perform better than students in the control condition. Results of an one-way ANOVA, however, indicated that there
were not statistically significant differences among conditions on recognition test total score, \( F (3, 224) = 1.45, p = .23 \).

Table 3-1: Comparative Means on Recognition Test Scores by Condition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Control ( n=63 )</th>
<th>Notetaking ( n=59 )</th>
<th>Self-questioning ( n=60 )</th>
<th>Summarizing ( n=46 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Recognition</td>
<td>9.71 (2.35)</td>
<td>9.76 (3.53)</td>
<td>8.73 (3.29)</td>
<td>9.89 (3.24)</td>
</tr>
<tr>
<td>Knowledge-level</td>
<td>1.62 (.83)</td>
<td>1.71 (.99)</td>
<td>1.53 (.97)</td>
<td>1.80 (.93)</td>
</tr>
<tr>
<td>Application-level</td>
<td>2.27 (.99)</td>
<td>2.02 (.90)</td>
<td>2.01 (.96)</td>
<td>2.04 (1.15)</td>
</tr>
<tr>
<td>ASE-level</td>
<td>6.00 (2.38)</td>
<td>6.19 (2.62)</td>
<td>5.27 (2.74)</td>
<td>6.15 (2.19)</td>
</tr>
</tbody>
</table>

**Knowledge-Level Items**

The second research question examined whether there were differences in students’ scores of knowledge-level recognition test items among the control, notetaking, self-questioning, and summarizing conditions. In knowledge-level item scores, we hypothesized that all the three comprehension strategy conditions would outperform the control condition but expected no difference among three strategy conditions themselves. (The means and standard deviations of scores on each level by strategy condition are presented in Table 3-1). Descriptive statistics results revealed students in the notetaking and summarizing strategy conditions scored higher on knowledge-level than students in the control condition.

An one-way ANOVA was then conducted to examine differences in knowledge-level item scores among strategy conditions. No significant differences were found among strategy conditions on knowledge-level items, \( F (3, 224) = .73, p = .45 \).
Application-Level Items

The third research question examined whether there are differences in students’ scores of application-level recognition test items among the control, notetaking, self-questioning, and summarizing conditions. We hypothesized that all the three comprehension strategy conditions would outperform the control condition but no differences among the three strategy conditions themselves. Interestingly, mean scores on application items were found to be highest in the control condition and lowest in the self-questioning condition (see Table 3-1).

An one-way ANOVA was then conducted to examine differences in application-level item scores among strategy conditions. No significant differences were found among conditions on application-level items, $F\,(3,224) =.89, p=.45$.

ASE-Level Items

The fourth research question explored differences in students’ scores of ASE-level (a combined level of analysis, synthesis, or evaluation) items. We hypothesized that using the notetaking strategy or using the summarizing strategy would result in greater performance on ASE-level items than the use of the self-questioning strategy or the control condition. As we expected, descriptively, those students in the notetaking and summarizing conditions scored higher than those in the self-questioning and control conditions (see Table 3-1).

The ANOVA conducted to examine outcomes on the ASE-level items indicated no statistically significant differences among conditions $F\,(3,224) =1.72, p=.16$. In contrast to our expectation, students in the self-questioning condition scored lowest among conditions.
Free Recall

The fifth research question examined whether there were differences in the number of idea units students generated during a free recall task among the control, notetaking, self-questioning, and summarizing conditions. It was hypothesized that students in the three comprehension strategy conditions would outperform students in the control condition on the number of idea units recalled. It was further hypothesized that students in the notetaking or summarizing conditions would outperformed those in the self-questioning condition, because notetaking and summarizing activities are similar to a practice of free recall. Means and standard deviations for students’ free recall by conditions are reported in Table 3-2. Overall, given students’ relatively low performance on the free recall across conditions, as we expected, students in the three comprehension strategy conditions recalled more total idea units than students in the control condition. Moreover, the notetaking condition, as we expected, outperformed the control and the other strategy conditions. The highest means of recalled total idea units was 5.97, which appeared in the notetaking condition; the lowest mean was 3.38, which appeared in the control condition (see Table 3-2).

Table 3-2: Comparative Means on Free recall by Condition.

<table>
<thead>
<tr>
<th>Free recall</th>
<th>Control</th>
<th>Notetaking</th>
<th>Self-questioning</th>
<th>Summarizing</th>
</tr>
</thead>
<tbody>
<tr>
<td>All participants $n=228$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total idea units</td>
<td>3.38 (4.04)</td>
<td>5.97 (7.13)</td>
<td>5.38 (4.95)</td>
<td>5.13 (4.66)</td>
</tr>
<tr>
<td>Ideas units from the text</td>
<td>2.21 (2.88)</td>
<td>3.64 (3.79)</td>
<td>3.17 (2.96)</td>
<td>2.80 (2.77)</td>
</tr>
<tr>
<td>Idea units from the text or diagram</td>
<td>1.17 (1.92)</td>
<td>2.32 (3.95)</td>
<td>2.22 (2.62)</td>
<td>2.33 (2.91)</td>
</tr>
</tbody>
</table>
Through Levene’s test procedure, the assumption of Homogeniety of Variances was violated in the total number of idea units $F(3, 224) = 3.109, p = .027$. Therefore, we conducted a Brown-Forsythe test instead of an one-way ANOVA to examine whether there were differences in “total idea units” among strategy conditions. The result indicated there was a significant strategy effect, $F(3, 185.17) = 2.70, p = .047$. Post hoc comparisons indicated that students in the notetaking condition performed significantly higher than students in the control condition on the “total idea units” ($p = .02$). As we expected, the notetaking condition performed the highest in free recall among conditions. However, those in the summarizing condition and self-questioning condition performed higher than the control condition but these differences were not statistically significant.

For analytical purposes, we split participants into three subgroups based on how many statistics courses they have taken. We hypothesized that for those who had taken statistics courses (1 course and 2+ course), no differences of recall among conditions would be expected; but for those who have never taken statistics courses (0 course), differences of recall were expected between the control and three comprehension strategy conditions. Means and standard deviations on free recall relative to statistics courses are reported in Table 3-3.

Within 0-course students, result of Brown-Forsythe statistic robust test indicated that there was a significant strategy effect on total idea units, $F(3, 60.08) = 3.22, p = .03$. Post-hoc comparisons indicated that students in the notetaking condition performed significantly higher than students in the control condition on the “total idea units” ($p = .01$). Further, we examined “idea units from the text” and “idea units from the text or diagram” in students’ recall, result of Brown-Forsythe test indicated a significant strategy effect on “idea units from the text” ($p = .01$), but no significant strategy effect on “idea units from the text or diagram” ($p = .17$). Results of post-hoc tests indicated that students in the three strategy conditions significantly outperformed
the control condition on “idea units from the text”. As we expected, results indicated that students
in the notetaking strategy condition recalled the most “total idea units”, “idea units from the text”
and “idea units from the text or diagram”. In contrast to our expectation, students in the
summarizing conditions did not always recall more idea units.

<table>
<thead>
<tr>
<th>Free recall</th>
<th>Condition</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>Notetaking</td>
<td>Self-questioning</td>
<td>Summarizing</td>
</tr>
<tr>
<td></td>
<td>0 course</td>
<td>1 course</td>
<td>1 course</td>
<td>2+ course</td>
</tr>
<tr>
<td></td>
<td>n=20</td>
<td>n=96</td>
<td>n=24</td>
<td>n=26</td>
</tr>
<tr>
<td>Total idea units</td>
<td>1.55 (2.72)</td>
<td>6.58 (8.82)</td>
<td>5.42 (4.35)</td>
<td>5.42 (5.19)</td>
</tr>
<tr>
<td>Idea units from</td>
<td>.95 (1.70)</td>
<td>4.04 (4.49)</td>
<td>3.29 (2.56)</td>
<td>2.88 (2.97)</td>
</tr>
<tr>
<td>the text</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Idea units from</td>
<td>.60 (1.23)</td>
<td>2.54 (4.95)</td>
<td>2.13 (2.47)</td>
<td>2.54 (3.30)</td>
</tr>
<tr>
<td>the text or diagram</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=30</td>
<td>n=25</td>
<td>n=28</td>
<td>n=17</td>
</tr>
<tr>
<td>Total idea units</td>
<td>3.80 (3.48)</td>
<td>6.12 (5.90)</td>
<td>5.29 (5.42)</td>
<td>4.65 (3.57)</td>
</tr>
<tr>
<td>Idea units from</td>
<td>2.40 (2.50)</td>
<td>3.60 (3.28)</td>
<td>2.96 (3.19)</td>
<td>2.59 (2.35)</td>
</tr>
<tr>
<td>the text</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Idea units from</td>
<td>1.40 (2.03)</td>
<td>2.52 (3.34)</td>
<td>2.32 (2.83)</td>
<td>2.06 (2.38)</td>
</tr>
<tr>
<td>the text or diagram</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n=13</td>
<td>n=8</td>
<td>n=8</td>
<td>n=3</td>
</tr>
<tr>
<td>Total idea units</td>
<td>5.23 (5.81)</td>
<td>3.5 (3.85)</td>
<td>5.63 (5.61)</td>
<td>5.33 (6.81)</td>
</tr>
<tr>
<td>Idea units from</td>
<td>3.69 (4.25)</td>
<td>2.50 (2.83)</td>
<td>3.50 (3.55)</td>
<td>3.33 (4.16)</td>
</tr>
<tr>
<td>the text</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Idea units from</td>
<td>1.54 (2.44)</td>
<td>1.00 (2.20)</td>
<td>2.13 (2.64)</td>
<td>2.00 (2.65)</td>
</tr>
<tr>
<td>the text or diagram</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Within 1-course students and 2-course students, results indicated that no differences of "total idea units", "idea units from the text" and "idea units from the text or diagram" in free recall among conditions. Within 1-course students, as we expected, the notetaking condition outperformed the other three conditions on "total idea units", "idea units from the text" and "idea units from the text or diagram" in free recall but with no significant differences. Interestingly, within 2-course students, neither notetaking nor summarizing condition performed the best among conditions.

**Time**

The sixth research question examined whether there were differences in the time that students spent in studying and test taking among conditions. In addition, we were also interested in whether the length of time students worked correlated to the recognition and recall learning outcomes. We hypothesized that the length of time students engaged in reading and taking tests would be positively correlated to learning outcomes such as recognition test total scores and the number of free recalled ideas. To prevent from extreme outliers affecting the correlations, we deleted 4 participants’ data, because they spent over 300 minutes on the task. The average time was 24.18 minutes for all participants. Therefore, the number of participants was changed from 228 to 224. Descriptive data and correlations of outcome variables are presented in Table 3-4.

An one-way ANOVA was conducted to examine whether there were differences in the time that students spent among conditions. Although results indicated no differences in time among strategy conditions, data revealed significant positive correlations between time and all outcome variables including “total idea units”, “idea units from the text”, “idea units from the text or diagram” in participants’ free recall and recognition test total scores. As would be expected,
results suggested that the more time students spent in reading with or without a comprehension strategy, the higher total recognition test scores they received and the more ideas they recalled.

Table 3-4: Descriptive Data and Pearson-r Correlations Among Outcome Variables.

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Total idea units</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Idea units from the text</td>
<td>.89**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Idea units from the text or diagram</td>
<td>.87**</td>
<td>.56**</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Total Recognition test score</td>
<td>.43**</td>
<td>.28**</td>
<td>.49**</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>5. Time (minutes)</td>
<td>.43**</td>
<td>.31**</td>
<td>.46**</td>
<td>.47**</td>
<td>-</td>
</tr>
<tr>
<td>M</td>
<td>4.98</td>
<td>2.97</td>
<td>2.01</td>
<td>9.48</td>
<td>24.18</td>
</tr>
<tr>
<td>SD</td>
<td>5.42</td>
<td>3.17</td>
<td>2.97</td>
<td>3.35</td>
<td>11.28</td>
</tr>
</tbody>
</table>

**p<.01.
Chapter 4

Discussion, Implications, and Future Research

Discussion

The primary purpose of this study was to evaluate the hypothesis that the comprehension strategies known to facilitate expository text comprehension were also effective in assisting students’ learning from complex diagrams. The strategy conditions we designed in this study included the use of notetaking (N), self-questioning (Q), and summarizing (S) strategies. These three strategies are known to facilitate expository text comprehension according to substantial empirical and theoretical evidence (e.g., Gajria, et al., 2007; Kiewra et al., 1989; 1991; NICHD, 2000). After we randomly assigned participants to one of the four experimental conditions, we examined students’ recognition and recall learning outcomes. We also considered the amount of time students spent on the learning and testing tasks. Overall, our results indicated a significant difference on the free recall task but no significant effect for assigned comprehension strategy condition on students’ recognition outcomes. Our results point to important conclusions, implications, and directions for future research.

Multiple-Choice Recognition Items

The first research questions examined students’ knowledge-level, application-level, and ASE- (analysis, synthesis, and evaluation) level items on a multiple-choice test. The overall means on the multiple-choice recognition test did not significantly differ by condition.
To further examine learners’ comprehension of different levels, the second research question addressed whether students’ comprehension, as measured by knowledge-level and application-level recognition test items differed based upon comprehension strategy condition. Theoretically, we hypothesized that N, Q, and S conditions would outperform the control condition (C) on these simple, lower-level items. Despite the prediction that all the three comprehension strategy conditions would outperform the control condition, analysis of significance tests failed to indicate a main effect for comprehension strategy on these items.

There may be three reasons why our prediction was not upheld. The first possibility was compensatory effects. We suspected that the control participants also used comprehension strategies such as summarization or self-questioning but without writing or typing down what they thought, questioned, or summarized. The comprehension strategies that the control participants used might have occurred in their mind, as a form of inner speech or inner dialogue (Janssen, 2002). For those in the N, Q, and S conditions, the demanding nature of generating and typing notes or questions may have distracted them or strained the limits of their working memory capacity (Atkinson & Shiffrin, 1968), thus reducing the effect of comprehension strategies.

Another explanation is that while participants were studying the text and diagram, they were not allowed to review the notes, questions, and summaries which they generated in the previous activity stage. Our study design did not allow student review of generated notes. In this study, our intent was to limit review to ensure all participants generated notes, questions, or summaries of the target diagram rather than the target text. Previous researchers have distinguished two functions of notetaking: encoding and external storage (Di Vesta & Gray, 1972; Kiewra et al, 1989; Kiewra et al, 1991). According to Kiewra and colleagues’ (1991) findings, encoding plus storage (notetaking with review) was superior to encoding (notetaking without
review) and superior to external storage (borrowing notes and review). We acknowledge that with no chance to review, participants might benefit less from comprehension strategies.

The third possibility is that participants in the C condition may have brought more effective personal strategies to the task other than notetaking, self-questioning, or summarization. Previous research had investigated various spontaneous study strategies employed by undergraduate learners. For example, Wade, Trathen, and Schraw (1990) concluded that there were six types of “studiers”: the Good strategy user, Information Organizer, Flexible Reader, Text Noter, Mental Integrator, and Memorizer. In other words, in the current study the control participants might be some types of these “studiers” and used their spontaneous strategies when they were reading. It would be meaningful for future research to explore how these spontaneous strategies function and whether they are effective.

In addition, we questioned whether learning from a diagram, as measured by ASE-level recognition test items differed based upon condition. Theoretically, we know that notetaking and summarizing facilitate higher-order thinking such as using more questions at a critical-thinking level (Hess, 2005; Jitendra et al., 1998), integrating materials (King, 1997), and improving inferences (King, 1997). Therefore, we expected that students using notetaking or summarizing would outperform students using self-questioning or free study (in the C condition) on ASE-level items. In the current study, however, our findings did not support a significant effect for notetaking and summarizing on ASE-level items. Previous research suggests that during reading, students need to activate relevant prior knowledge to enhance questioning, and restructure the schemata. The consequence of such activation and restructuring is better comprehension (Gajria et al., 2007; Janssen, 2002). Accordingly, learners with more prior knowledge may benefit more from self-questioning than learners with limited prior knowledge. For learners with limited knowledge, self-questioning may be not the optimal strategy for learning from diagrams.
Free Recall Measure

Our fifth research question examined the number of idea units students recalled varied based upon strategy condition. We hypothesized that the learners in the three comprehension strategy conditions, N, Q, and S would outperform the control condition in free recall. Overall, given students’ relatively low performance in free recall across conditions, analysis of total ideas indicated that there was a significant strategy effect and students in the notetaking condition significantly outperformed the control condition. The current findings were consistent with previous research that indicated that comprehension strategies facilitate students’ knowledge recall (Cordero-Ponce, 2000; Elosua et al., 2002; Kiewra et al., 1991; Kobayashi, 2005; Rich & Shepherd, 1993). The current findings reveal that although the effectiveness of self-questioning and summarizing strategies is still unclear, notetaking is an effective strategy in assisting students’ knowledge recall from an expository text in conjunction with a complex diagram.

Time

The last research question explored whether the amount of time that students spent differed by condition. Results indicated no significant differences in time spent among conditions. However, results also indicated positive correlations between the amount of time and two learning outcomes. Both recognition test scores and free recall idea units were positively correlated with the amount of time that participants spent. The longer time that students spent engaged on learning tasks, the more idea units they recalled. The correlations between time and outcome measures were moderate and positive across conditions. The current findings suggested that the more time in reading with N, Q, S strategy or their own strategies, the better students’
understanding from complex diagrams. Shanahan (2005) noted that comprehension strategies slow down readers:

To use strategies well, the student has to be reflective and purposeful; instead of trying to do something quickly without paying attention, strategies slow the reader down and focus his or her attention according to the demands of purposes and needs. When someone wants to understand and remember a text very well, he or she should preview the text carefully to form a clear idea of what it might be about; think about what is already known about a topic or make predictions about what information will be presented; stop along the way during reading, and ask questions about what the text says (and try to answer these self-posed questions); and summarize the text occasionally to make sure it is being understood. None of these actions speeds a reader along. (p. 29)

We view it is necessary to explore this relationship between the amount of time students used and strategy effectiveness. Future research should continue to explore the optimal time learners should engage with instructional materials for the effective use of comprehension strategies.

Implications

To evaluate the hypothesis that the comprehension strategies known to facilitate expository text comprehension were also effective in assisting students’ learning from diagrams, the current research reported a significant strategy effect only for notetaking on knowledge recall, but no significant strategy effect for self-questioning and summarizing on knowledge recall and recognition learning outcomes. The current research suggested that learners benefit from a notetaking strategy when learning from a complex diagram and enhance their knowledge recall. However, whether a comprehension strategy can assist students’ recognition or comprehension of diagrams, findings are still inclusive. The results of our study, however, provide meaningful contributions to our collective understanding of learning from diagrams with comprehension
strategies. There are two major implications from the current study for comprehension strategy use when learning from expository text in conjunction with complex diagrams.

First, on a practical level, the current study shows that college students can be successfully prompted to use a notetaking strategy to facilitate their knowledge recall from expository text in conjunction with complex diagram. In addition, although not significantly, students using self-questioning or summarizing strategies did recall more knowledge than those in the control condition. Students who face difficulty comprehending or have limited prior knowledge are less likely to engage in spontaneous comprehension strategies such as notetaking or summarizing on their own during learning (Bransford et al., 1982; Brown & Day, 1983). Perhaps college instructors should be advised to provide this sort of support and guidance to their students by demonstration of notetaking when teaching complex diagrams, or by teaching them to use a notetaking strategy.

Second, we found that whether on knowledge-level, application-level, or ASE-level items of recognition test, although not significantly, students using self-questioning performed worse than those in the other three conditions including the control condition. However, we acknowledge that the manner in which we required participants in the self-questioning condition to engage in actions like generating questions or typing questions was potentially distracting, resulting in more concentration on information they did understand rather than what they could learn from the diagram. This may be the reason why our findings illustrate limited effectiveness of the self-questioning strategy for recognition test performance. Therefore, it is also suggested that college students may not be encouraged to ask themselves too many noncrucial questions when learning from complex diagrams, especially those with low-prior knowledge. As students are not able to generate “good” questions, noncrucial questions or content-irrelevant questions may distract students’ attention and strain working memory capacity (Atkinson, & Shiffrin, 1968). College students with limited knowledge are prompted to freely ask themselves questions or
make summaries when studying a complex diagram, as a result, they may have better knowledge recall from the materials.

**Future Research**

We envision five major areas for future research. The first and second major areas focus on the instructional features of comprehension strategies. The third suggests alternative outcome measures, such as transfer tests or retention tests. The fourth area of suggested future research is to explore how types or qualities of generated notes, questions, and summaries affect comprehension. The last future research direction focuses on students’ characteristics.

First, the current findings suggest students using a notetaking strategy recall significantly more ideas than students in the control condition but there was no significant effects for strategy condition on recognition tests. In other words, findings from this study suggest N, Q, and S are not more effective strategies for college students on performing recognition test when reading complex diagrams in conjunction with expository text. One possibility may be that students are not able to effectively generate “helpful” notes when using notetaking (Bretzing & Kulhavy, 1981; Kiewra & Fletcher, 1984); are not able to elaborative “why” questions when self-questioning (Rosenshine, Meister & Chapman, 1996; Wong, 1985); and are not able to generate “good” summaries when summarizing (Brown & Day, 1983). It may be that students need more instruction and practice to effectively engage in the strategies. The majority of previous researchers concluded that when students practice or receive training in how to take notes by producing such specified forms as a concept map, a matrix, or an outline of the representations (Bretzing & Kulhavy, 1981; Katayama & Robinson, 2000; Kiewra & Fletcher, 1984; Robinson et al., 2006); in how to generate questions (Rosenshine, Meister & Chapman, 1996; Wong, 1985; King, 1992); and in how to make a summary by adapting main idea strategy, metacognitive
strategy instruction, or reciprocal teaching (Parlincsar & Brown, 1984; Cordero-Ponce, 2000; Jitendra, et al., 2000; Ferguson, 2001; Weedman, 2003), students’ use of comprehension strategies results in improved comprehension and recall during or after reading (Kobayashi, 2005; Kobayashi, 2006a; 2006b). Previous findings also suggest strategy instruction that includes the value, purpose, and self-monitoring of comprehension strategy is more effective in increasing text comprehension than using a strategy alone (Ferguson, 2001; King, 1992). Obviously, notetaking, self-questioning, or summarizing a verbal presentation such as a text passage differs from these actions of a visual representation such as a diagram. Future research might test the instructional effectiveness of strategies to comprehend complex diagrams instead of expository texts.

The second area of future research is effectiveness of multiple strategies on learning from diagrams. The majority of previous research has concluded that students’ expository comprehension benefits most from using multiple strategies or receiving combined strategy training (Lubliner, 2002; Philbrick, 2003; Weedman, 2003). More recently, results of a meta-analysis indicated larger effects for instruction in multiple comprehension strategies when compared to a single strategy for students with learning disabilities (Gajria, et al., 2007). Future research might examine whether comprehension of diagrams can be better enhanced by integrating use of various cognitive strategies.

Third, previously, Malone and Mastropieri (1992) demonstrated positive results for summarizing instruction on immediate posttest, near transfer, and far transfer measures. Recently, additional research indicated that participants in summarizing groups outperformed a no-strategy group on recalling more ideas of expository text both at the immediate test and at delayed test (Cordero-Ponce, 2000), and maintained summarizing usage on the near transfer measure but not far transfer measure (Jitendra, Hoppes & Xin, 2000). In the current study, however, our outcome measures did not include a delayed or a transfer test. Further research might examine the effectiveness of comprehension strategies on delayed or transfer tasks.
Fourth, we also collected student-generated notes, questions, and summaries in the
current study but we did not score them because they were not directly relevant to our research
questions. Future research might explore types or qualities of student-generated notes, questions,
and summaries when learning from diagrams to inform our understanding of comprehension
strategies.

A fifth area of future research would be to consider student characteristics such as age or
ability level. College students or adult learners likely have developed many effective strategies
for use when they face challenging comprehension situations. Younger children, such as
elementary or junior high school students, however, may have less developed strategies. Cheung
(1995) also suggested that especially below-average students benefited from explicit strategy
training. Future research might examine the effectiveness of comprehension strategies with
younger children or below-average learners’ comprehension of expository texts and diagrams.

There are a few limitations of the study we would like to note. First, as previously
mentioned, the participants were predominately female college students who are already
successful, in terms of university enrollments. Hence the findings may be limited in their
generalizability to other populations such as younger children or below-average learners. A
second limitation of the current study is the sole reliance on the use of free recall and the
recognition test measures as criteria used to gather learning outcomes.

In conclusion, the current study contributes to our understanding of comprehension
strategies in several ways. First, the current study extends the existing knowledge base of
comprehension strategies from learning an expository text and explores the use of comprehension
strategies for learning from a complex diagram. A large body of previous comprehension strategy
research has addressed comprehension of expository texts, however, few have studied
comprehension of diagrams. In this study we explored the effectiveness of comprehension
strategies with a statistical expository text in conjunction with a complex diagram. Second, the
current work also contributes to our existing knowledge base of comprehension strategy usage in an on-line environment. Therefore, this study tests the effectiveness of comprehension strategies in an ecologically-valid learning setting. Additionally, the study also introduces critical variables to successfully comprehend diagrams. Both learners’ prior content exposure and the amount of time learners invested in reading were discussed in the study.

It is important to highlight that comprehension strategy effects on learning from texts and from diagrams are quite complex, and may be quite different. Readers daily face challenges with the comprehension of complex learning materials. Future research should continue to explore the benefits of comprehension strategy use in these complex learning environments.
Bibliography


Won, W. (2001, November). *The relative effectiveness of structured questions and summarizing on near and far Transfer tasks.* Papers Presented at the National Convention of the Association for Educational Communications and Technology, Atlanta, GA.


## Published Experimental Studies of Notetaking Strategy (1972-2007)

<table>
<thead>
<tr>
<th>Studies</th>
<th>Participants</th>
<th>Reading material</th>
<th>Experiment/conditions</th>
<th>Main findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Di Vesta &amp; Gray</td>
<td>Students; N = 120</td>
<td>Different topics; each topic contains 500 words.</td>
<td>20 experimental conditions. Ss listened to a set of 3 5-min passages with 4 orthogonally crossed variables: position of the criterion passage on an imaginary scientific system in the set, note taking while listening, rehearsal immediately after listening, and testing.</td>
<td>There were more words generated and higher multiple-choice test scores when the study interval was used for review than for other activities. The number of ideas recalled was favorably influenced by note taking, rehearsal, and testing. There were no significant effects due to position of the passage in the set. Significant correlations were obtained between performance and the individual difference variables of anxiety and tolerance of ambiguity.</td>
</tr>
<tr>
<td>Kiewra (1989)</td>
<td>Undergraduates; Exp1, N=135, Exp2, N=94</td>
<td>Lecture and texts</td>
<td>Group1: take notes without reviewing (encoding only); Group2: take notes and review (encoding-plus-storage); Group3: borrow notes and review (external-storage). 3 forms of note taking were used: conventional; note taking on skeletal; and matrix frameworks.</td>
<td>Conducted 2 studies that investigated the (1) encoding, (2) encoding-plus-storage, and (3) external-storage functions of note taking. In both Exp 1, which used 135 Ss, and involved lecture learning, and Exp2, which used 94 Ss, and involved text learning, an advantage was found for the encoding-plus-storage function on tests involving factual-recall and recognition performance, but not on tests measuring higher-order performance. When test material was used, there was some advantage for conventional notes, and a clear advantage for not taking notes at all.</td>
</tr>
<tr>
<td>Kiewra, et al (1991)</td>
<td>Undergraduates; Exp1, N=135, Exp2, N=94</td>
<td>Lecture and texts</td>
<td>Group1: take notes without reviewing (encoding only); Group2: take notes and review (encoding-plus-storage);</td>
<td>Results pertaining to note-taking functions indicated that encoding plus storage was superior to encoding and to external storage for recall performance, and superior to encoding for</td>
</tr>
</tbody>
</table>
Group3: borrow notes and review (external-storage). Three forms of note taking were used: conventional, note taking on skeletal; and matrix frameworks. 

Van Meter, P., Yokoi, & Pressley (1994)  
Undergraduates; N=252  
Lectures  
Individual interview  
Group Interviews  
Lecture style variables: notes differed in difficulty as a function of lecturer. Students reported they have more difficulty in taking notes if the lecture are disorganized or rapid paced. Student variables: connecting Prior knowledge (Pepper & Mayer, 1986), or students; developmental maturity  
Course content variables:  
- Post-class processing of notes

Robinson & Kiewra (1995)  
Undergraduates; N= 42  
A chapter taken from abnormal behavior in textbook (6500 words)  
Exp I: graphic organizer group; outline group; text only group.  
Exp II: text only group; text plus outlines group; text plus graphic organizers.  
When given enough time, students studying graphic organizers learned more hierarchical and coordinate relations, and they were more successful in applying that knowledge and in writing integrated essays than students studying outlines or text alone.

Faber, Morris, & Lieberman (2000)  
Grade 9; N=155  
World culture  
Note taking instruction; control group; (low interest v. s.high interest)  
The note taking instruction group and control group scored significantly different. The difference arose mainly from the low-interest passage test. (not ability level, or passage type)

Katayama & Robinson (2000)  
Undergraduates; N=117  
Lecture  
Graphic organizers group; Outlines group; Partial notes v. s. complete notes  
Measure: factual test v. s. application test  
Results with 117 undergraduates show no effect on the factual test for either study notes or amount of information. But on the application test, graphic organizers were better than outlines and partial notes were better than complete notes.

Porte (2001)  
Grade 9 and 10 students with emotional and Grade 9 English  
Note taking strategy and review  
A new note taking strategy emphasizes manipulating and organizing information rather than writing it.
<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Materials</th>
<th>Task</th>
<th>Method</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trafton &amp; Trickett (2001)</td>
<td>Undergraduate; N=90.</td>
<td>Online notepad 5 problem solving questions. correlation 2. experimental</td>
<td>Problem solving</td>
<td>Using the notepad helped participants solve the problems more accurately; The benefits of using the notepad persisted after participants had stopped using it; and Participants who used the notepad for problem-solving and self-explanation learned more, regardless of the type of notepad interface that was provided.</td>
<td></td>
</tr>
<tr>
<td>Katayama &amp; Grooks (2003)</td>
<td>Undergraduates; N= 67.</td>
<td>2 electronic notes conditions (complete vs. partial) and two testing conditions (immediate vs. delayed) on three types of tests (fact, structure, and application).</td>
<td>A blended chapter length-text (approx 3500 words)</td>
<td>The results did show a significant main effect for notes conditions (partial &gt; control) on the application test and for testing condition. Students tested immediately performed significantly better than those in the delayed condition on all three outcome measures. Results also show significant interactions between Notes Condition x Testing Condition on the structure and application tests.</td>
<td></td>
</tr>
<tr>
<td>Peverly; Brobst; Graham; Shaw (2003)</td>
<td>College students; N=88</td>
<td>Notes group; No-notes group. (College adults are not good at self-regulation: A study on the relationship of self-regulation, note taking, and test taking)</td>
<td>History text</td>
<td>Research using simple materials has found that adults are skilled self-regulators. Research using difficult materials has found the opposite. Results imply that test performance is more related to note taking and background knowledge than to self-regulation.</td>
<td></td>
</tr>
<tr>
<td>Katayama, Shambaugh, &amp; Edmonds (2005)</td>
<td>Undergraduates; N= 47</td>
<td>Copy-and-paste note taking group; typed note taking group.</td>
<td>2 electronic chapter-length texts</td>
<td>Keying in notes leads to higher retention of knowledge transfer than copying and pasting notes after a 1-week delay.</td>
<td></td>
</tr>
<tr>
<td>Igo (2005)</td>
<td>Undergraduates</td>
<td>7 words group; 14 to 21 words group; unrestricted group.</td>
<td>Not mention</td>
<td>High achieving students in the 14- and 21-word restriction groups were likely to count words and scroll through the cells of their note-taking charts, which perhaps distracted them from learning the text ideas. Students in the seven-word or unrestricted groups tended not to count words or scroll through the cells, they were not distracted, and consequently, learning was not deterred.</td>
<td></td>
</tr>
<tr>
<td>Igo, Bruning, &amp; McCrudden (2005)</td>
<td>Undergraduate; N= 71</td>
<td>graphic organizers group; outlines group;</td>
<td>A text passage describing three</td>
<td>Note taking can be enhanced by (a) using graphic organizers rather than outlines, (b) providing cues</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Grade</td>
<td>N</td>
<td>Learning style</td>
<td>Instruction</td>
<td>Measures</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>---</td>
<td>---------------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>Brown (2005)</td>
<td>Grade 7, N=66</td>
<td>A core history text</td>
<td>Questionnaire, tapes.</td>
<td>The results indicate that even though students verbalized their meta-cognitive understandings with a fair degree of sophistication, they were far less proficient in their self-regulated note-taking practices.</td>
<td></td>
</tr>
<tr>
<td>Trasborg (2005)</td>
<td>Grade 9 or 10; N=150</td>
<td>Global studies class: expository text</td>
<td>Pretest (Nelson Denny, Form H), 5 instruction sessions and a posttest (Nelson Denny, Forms G): direct explanation only condition (DE-only); direct explanation and note taking condition (DE+NT); control condition.</td>
<td>The posttest results show that direct explanation and note taking condition DE+NT was a significantly more effective training method than DE-only for improving the reading comprehension of high school students. Students in DE+NT and DE-only training groups significantly outperformed the controls that were given no instruction.</td>
<td></td>
</tr>
<tr>
<td>Dolly (2006)</td>
<td>Undergraduates; less skilled readers; N=48</td>
<td>Lecture</td>
<td>Authentic study; Note taking log 5 times During 15 weeks</td>
<td>The four case study students reported and demonstrated increased metacognitive behavior. Students increased their questioning behavior, and positive perceptions of themselves as readers.</td>
<td></td>
</tr>
<tr>
<td>Robinson and others (2006)</td>
<td>Undergraduates; Exp I: N= 114 Exp II: N= 120 Exp III: N= 110 Exp IV: N= 58</td>
<td>12 chapters of the textbook “Human Learning”</td>
<td>Exp I to III: partial notes group; complete notes group. Exp IV: groups 1-6 served as the treatment group; groups 7-12 were the control group. Treatment group used the Web-based GO activity for the first three units, and used the paper-based notes for the last three units.</td>
<td>The partial task led to increased overall examination performance in all experiments. In all experiments, GO note taking increased. The increases were greatest when the authors presented the partial task in a computer environment with a timed, forced-choice task. Implications for using the partial GO task in the classroom, as well as future note-taking research directions are discussed.</td>
<td></td>
</tr>
<tr>
<td>Peverly et al. (2007)</td>
<td>Undergraduates; N=85</td>
<td>Lecture</td>
<td>Transcription fluency verbal working memory ability of identify main idea</td>
<td>Transcription fluency is the only predictor of quality of notes, and quality of notes was the only predictor of test performance.</td>
<td></td>
</tr>
</tbody>
</table>

---

D. from Dissertation Abstracts International Database.
## Appendix B

### Published Experimental Studies of Self-Questioning Strategy (1982-2002)

<table>
<thead>
<tr>
<th>Studies</th>
<th>Participants</th>
<th>Reading material</th>
<th>Experiment/conditions</th>
<th>Main findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wong &amp; Jones (1982)</td>
<td>Grade 8 and 9 learning disabled (LD) N=60; Grade 6 normally achieving N=60</td>
<td>Class material</td>
<td>5-step self-questioning training</td>
<td>Results show that training substantially increased LD Ss' awareness of important textual units, as well as their ability to formulate questions involving those units. Training also facilitated their comprehension performance. However, training did not substantially increase normally achieving Ss' metacomprehension or comprehension performance.</td>
</tr>
<tr>
<td>Palincsar &amp; Brown (1984)</td>
<td>Exp I: Grade 7; poor readers; N=24; normal readers N=13. Exp II: N=21</td>
<td>Lectures</td>
<td>Reciprocal teaching summarizing, questioning, clarifying, and predicting</td>
<td>Reciprocal teaching with the adult guiding the Ss led to improvement in the quality of summaries and questions, gains in criterion tests, reliable maintenance over time, task transfer, and generalization to the classroom setting. Exp 2 replicated many of these results, and was useful in helping to understand the underlying cognitive mechanisms involved in reading and studying.</td>
</tr>
<tr>
<td>Scruggs and Mastropieri (1985)</td>
<td>10-13 yr old gifted students N= 21; nongifted students N= 23</td>
<td>Lecture</td>
<td>Gifted group; nongifted group.</td>
<td>Gifted Students outperformed their age peers in recall and strategy use for both meaningful and nonmeaningful word pairs, and they appeared to benefit greatly from mediational strategy use. Gifted-students who used elaborative interrogation strategies learned more than non-gifted students.</td>
</tr>
<tr>
<td>Forsythe (1986)</td>
<td>Grade 5; below average students; N=37</td>
<td>Social studies</td>
<td>Experimental group (SQ3R strategy): surveying, questioning, reading, reviewing, reciting, organizing and reading skills, skimming, note taking, and reading charts, maps and diagrams</td>
<td>The experimental sample showed a significant increase in student achievement in the content area of social studies as shown on chapter tests given covering the work.</td>
</tr>
<tr>
<td>Study</td>
<td>Grade/Group Description</td>
<td>Type of Study</td>
<td>Description</td>
<td>Findings</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------------</td>
<td>---------------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>MacGregor (1988)</td>
<td>Grade 3 average/good readers; N=48</td>
<td>Lecture</td>
<td>Experimental group: computer-mediated text system (CTS) and were trained to generate questions for clarification and/or focus of attention; Control group: no training.</td>
<td>CTS would be effective for teaching Ss to generate questions to improve reading performance. Average readers made significantly greater gains than good readers in comprehension and vocabulary knowledge; Good readers made greater gains in ability to predict vocabulary performance.</td>
</tr>
</tbody>
</table>
| King (1989)      | Undergraduates; N=32    | Lecture | 1) independent self-questioning (S-Q)  
2) self- and peer-questioning in small cooperative groups (SCGs),  
3) review in SCGs,  
4) independent review. | Both of the S-Q strategies significantly improved lecture comprehension over time. Significant differences were found among the strategies with the self/peer questioners as well as the self-questioners showing posttreatment comprehension superior to that of Ss using either of the review strategies. No significant change was found for Ss’ feelings of control. |
| Shiang & McDaniel (1991) | Undergraduates; N=33 | Lecture | Investigate thinking progress under 3 conditions:  
external higher order questions;  
external low order questions;  
self-generated question. | Question conditions were unrelated to the quality of explanations of the complex situation constructed by Ss. Note-taking was related to the quality of the explanations. Questions and prompts to generate questions may elicit overt compliance without necessarily motivating Ss to interrogate the material. |
| King (1991)      | Grade 9; N=56           | Lecture | Self-questioning only; self-questioning with reciprocal peer-questioning; review condition; control group. | On postpractice and 10-day maintenance tests, Ss in the self-questioning with reciprocal peer-questioning and the self-questioning-only strategy groups showed lecture comprehension superior to that of participants in both the discussion review and control groups. |
| King (1992)      | Undergraduates; N=56    | Lecture | Self-questioning condition; summarizing condition; note-taking review condition. | Self-questioners performed better than summarizers and significantly better than note-reviewers |
| Penkingcarm (1992) | Grade 11-12             | Expository texts | SQ3R questioning group: generating questions with SQ3R; Questioning group: generating questions without SQ3R; Control group: non questioning. | Both questioning groups scored higher on reading comprehension test than the control group. |
| Scardamalia &    | Grade 5 and 6; “Wildlife of Exp I: | Expository texts | Knowledge-based questions formulated in advance of | |


<table>
<thead>
<tr>
<th>Study</th>
<th>Grade</th>
<th>Sample Description</th>
<th>Interventions</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bereiter (1992)</td>
<td>2 classes</td>
<td>Jeopardy”</td>
<td>Knowledge-based condition; text-based condition. Exp I: Basic questions; Wonderment questions.</td>
<td>Instruction were found to be of a higher order than text-based questions produced after exposure to text materials.</td>
</tr>
<tr>
<td>King &amp; Rosenshine</td>
<td>Grade 5 N=34</td>
<td>Science material about “tide pools”</td>
<td>Training in use of question stems (explain why … is important?); Training in use of signal words (what, where, why?); Unguided questioning.</td>
<td>The group that was trained in using question stems outperformed the other two group</td>
</tr>
<tr>
<td>Rich &amp; Shepherd</td>
<td>Adult poor readers, in adult education program</td>
<td>Expository texts on science and social studies topics.</td>
<td>Self-questioning and summarizing; Self questioning only; Two control groups: no reading instruction;</td>
<td>Both self-questioning groups outperformed control groups on question task. The self-questioning and summarizing group outperformed the two control groups on free recall task.</td>
</tr>
<tr>
<td>Vittayarungrangsri</td>
<td>Higher education students</td>
<td>Expository texts</td>
<td>Student questions group: training in question generation by students; teacher-questions group: teachers questions about text</td>
<td>Student questions group outperformed the teacher-questions group on reading comprehension, and on number of questions generated by students.</td>
</tr>
<tr>
<td>King (1994)</td>
<td>Grade 4-5 N=48</td>
<td>Science material about “the system of the body”</td>
<td>Lesson-based questioning group: training of explaining; lesson-based questioning and experience-based questioning; Control group: unguided questioning.</td>
<td>Experimental groups outperformed the control group. Indications that experience-based questions are superior to lesson-based questions. When children use questions that guide them to connect ideas with a lesson together or connect the lesson to their prior knowledge, they engage in complex knowledge construction that, in turn, enhances learning.</td>
</tr>
<tr>
<td>Cheung (1995)</td>
<td>Grade 9 below-average students</td>
<td>Short and long text expository passages.</td>
<td>Self-questioning group: training in self-questioning; Control group: no training.</td>
<td>Self-questioning group scored higher on comprehension and inference test. Especially below-average students benefited from the training.</td>
</tr>
<tr>
<td>Westera &amp; Moore</td>
<td>Grade 8, poor comprehenders</td>
<td>Books, expository and narrative articles</td>
<td>Extended reciprocal teaching program (12–16 sessions); Short reciprocal teaching program (6–8 sessions); No-treatment comparison group.</td>
<td>Students in extended program showed significant gains in comprehension. No significant differences between the short program and the comparison group were found.</td>
</tr>
<tr>
<td>Brown, Pressley, Van Meter &amp;</td>
<td>Grade 2; low achievers</td>
<td>Illustrated stories or fables</td>
<td>Experimental group: Students Achieving Independent Learning</td>
<td>SAIL-group was more interpretive in story retellings, recalled more information, produced more reader-based</td>
</tr>
<tr>
<td>Author</td>
<td>Subjects</td>
<td>Text Type</td>
<td>Intervention</td>
<td>Outcome</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Schuder (1996)</td>
<td>Undergraduates</td>
<td>Expository texts</td>
<td>Program (SAIL), training in various reading strategies;</td>
<td>responses and scored higher on reading comprehension and word skills, than comparison group.</td>
</tr>
<tr>
<td></td>
<td>N=86</td>
<td></td>
<td>Comparison group: conventional reading instruction.</td>
<td></td>
</tr>
<tr>
<td>El-Koumy (1996)</td>
<td>Undergraduates</td>
<td>Expository texts</td>
<td>Reciprocal teaching group: reciprocal teaching of strategies;</td>
<td>Reciprocal teaching group scored higher on reading comprehension than teacher questioning group, which in turn scored higher than student-generated-question group.</td>
</tr>
<tr>
<td></td>
<td>N=86</td>
<td></td>
<td>Student-generated questions group;</td>
<td>Reciprocal teaching&gt; teacher questioning &gt; student-generated-question</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Teacher-provided questions group;</td>
<td></td>
</tr>
<tr>
<td>Neber (1996)</td>
<td>Grade 7</td>
<td>History tests, “the age of discovery”</td>
<td>Causal questions group: training in questioning for causes of historical events and people’s motives;</td>
<td>The causal questions group reached higher knowledge levels than the factual questions group.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Factual questions group: training in questioning for historical facts.</td>
<td></td>
</tr>
<tr>
<td>Trabasso &amp; Magliano (1996)</td>
<td>Undergraduates; N=8</td>
<td>4 hierarchical and 4 sequential versions of 8 stories</td>
<td>Verbal protocol and think-aloud data were obtained.</td>
<td>Coherent understanding is achieved by keeping available via maintenance or retrieval higher order goal information that explains a number of actions and outcomes of a plan. Readers predominantly attempt to explain why events, states and actions occur in a story</td>
</tr>
<tr>
<td>Ezell, Hunsicker &amp; Quinque (1997)</td>
<td>Grade 4</td>
<td>Stories, taken from basal readers</td>
<td>Peer-assisted procedure group: (students asking and answering questions); Teacher-assisted procedure group: (individual questioning, class discussion).</td>
<td>There was no significant difference between the groups in reading comprehension and question abilities.</td>
</tr>
<tr>
<td>Glaubman, Ofir &amp; Glaubman (1997)</td>
<td>Kindergartens N=93</td>
<td>Stories from young children</td>
<td>Metacognitive group: training to ask metacognitive questions; Questioning groups: active processing approach; Control group: Traditional instruction.</td>
<td>Both questioning groups showed higher gains on story comprehension than control group. Metacognitive group was superior in questioning and self-directed learning.</td>
</tr>
<tr>
<td>King (1997)</td>
<td>Grade 7</td>
<td>Science material</td>
<td>Sequenced-inquiry group: training in tutoring with sequenced questioning;</td>
<td>Sequenced-inquiry group outperformed the other groups on comprehension and integration of material, inferencing and knowledge recall.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unsequenced-inquiry group: training in</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Participants</td>
<td>Context</td>
<td>Treatment 1</td>
<td>Treatment 2</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Kitajima (1997)</td>
<td>Undergraduates</td>
<td>Stories about experiences in Japan</td>
<td>Experimental group: strategy training, focused on monitoring referential ties in texts. First step is forming questions; Control group: comprehension exercises.</td>
<td>Positive effects on students’ comprehension and ability to identify referents in texts, compared to control group.</td>
</tr>
<tr>
<td>Klingner, Vaughn &amp; Schumm (1998)</td>
<td>Grade 4, N=141</td>
<td>Social studies texts from a history textbook</td>
<td>Reading strategies group: reading strategies in small student-led groups (e.g. predicting test-questions after reading); Control group: researcher-led instruction (without strategies).</td>
<td>Reading strategies group made greater gains in reading comprehension than control group, and equal gains in content knowledge.</td>
</tr>
<tr>
<td>King, Staflieri &amp; Adelgais (1998)</td>
<td>Grade 7</td>
<td>Science material concerning human body</td>
<td>Peer tutoring group (no questioning); Sequenced questioning group; Unsequenced questioning and explanation group; Explaining-only group;</td>
<td>Both question-groups scored better on knowledge tests than explanation only group. No significant differences between sequenced and unsequenced questioning were found.</td>
</tr>
<tr>
<td>Niemczyk, Mary Catherine (2002)</td>
<td>High school; N= 107</td>
<td>Civil rights movement in the United States on learning from a multimedia database</td>
<td>Adjunct questions condition; student self-generated questions condition; note taking condition.</td>
<td>Students in the note-taking treatment condition performed significantly better than students working in the self-generated question treatment condition. Results did not yield significant differences in posttest scores among the treatment conditions.</td>
</tr>
</tbody>
</table>

D: from Dissertation Abstracts International Database.
## Published Experimental Studies of Summarization Strategy (1978-2005)

<table>
<thead>
<tr>
<th>Studies</th>
<th>Participants</th>
<th>Reading material</th>
<th>Experiment/conditions</th>
<th>Main findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doctorow, Wittrock &amp; Marks (1978)</td>
<td>Grade 6 N=488</td>
<td>Stories</td>
<td>Summary group; Summary and headings given group; Control group.</td>
<td>The facilitation of generative processes by the insertion of paragraph headings and instructions to generate sentences about story paragraphs during encoding produced the greatest comprehension, followed in turn by instructions to generate sentences, the insertion of paragraph headings, and then by reading the same stories without generative instructions or paragraph headings. The combination of inserted paragraph headings and instructions to generate sentences about paragraphs approximately doubled comprehension and recall in each experiment.</td>
</tr>
<tr>
<td>Brown &amp; Day (1983)</td>
<td>Grade 5, 7, 10; undergraduates; grade students. N=54</td>
<td>Two expository texts: &quot;Desert&quot; and &quot;Noise&quot;.</td>
<td>Ss were told they can do anything such as making summaries, taking notes, writing drafts before they write a 60-word summary.</td>
<td>Older high school and college Ss were able to use sophisticated condensation rules, such as invention and integration, in contrast to the 5th and 7th graders who relied on a more simple copy-delete strategy. Graduate English students who had taught rhetoric courses outperformed freshman college students in their ability to combine information across paragraphs and in their propensity to provide a synopsis in their own words. College students experienced particular problems with critical reading and effective studying.</td>
</tr>
<tr>
<td>Wittrock &amp; Kelly (1984)</td>
<td>Adult soldiers with low reading ability</td>
<td>Several reading passages</td>
<td>Ss were taught basic strategies including imagery, verbal, summarizing, metacognitive strategies.</td>
<td>The soldiers in each of he experimental groups showed statistically significantly greater improvement in reading comprehension than did the soldiers in the control groups.</td>
</tr>
<tr>
<td>Grave (1986)</td>
<td>Grade 5-8;</td>
<td>Several reading</td>
<td>MTC plus DI condition:</td>
<td>Results indicate that both MTC plus DI and DI alone were</td>
</tr>
<tr>
<td>Study</td>
<td>Grade</td>
<td>Learning Status</td>
<td>Material</td>
<td>Training Conditions</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>-------</td>
<td>-----------------</td>
<td>---------------------</td>
<td>--------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Armbruster, Anderson &amp; Ostertag (1987)</td>
<td>Grade 5</td>
<td>N= 82</td>
<td>Social studies material</td>
<td>Structure training group (which received direct instruction in recognizing); summarizing group traditional training group (which read and discussed answers to questions).</td>
</tr>
<tr>
<td>Grave &amp; Levin (1989)</td>
<td>Grades 5-8; learning disabled; N = 30</td>
<td>Several reading passages</td>
<td>Control condition; monitoring condition; mnemonic condition.</td>
<td></td>
</tr>
<tr>
<td>Ellis &amp; Graves (1990)</td>
<td>Grades 5-7; learning disabled; N = 68</td>
<td>Narrative passages; expository passages</td>
<td>Paraphrasing cognitive strategy condition; repeated readings condition; paraphrasing plus repeated readings condition; control condition.</td>
<td></td>
</tr>
<tr>
<td>Gajria &amp; Salvia (1992)</td>
<td>Grade 6-9; learning disabled N=30.</td>
<td>Expository prose</td>
<td>Experimental group: Ss were trained to use 5 rules of summarization: reduce lists, select topic sentences, construct topic, delete redundancies, and delete unimportant information. Control group: Ss had no training.</td>
<td></td>
</tr>
<tr>
<td>Wittrock &amp; Alesandrini (1990)</td>
<td>Undergraduates N= 59</td>
<td>Rachel Carson's book <em>The Sea Around Us.</em></td>
<td>Group1: generate Summaries; Group 2: generate Analogies; Group 3 (control): read Text.</td>
<td></td>
</tr>
<tr>
<td>Malone &amp;</td>
<td>Middle-school</td>
<td>Expository passages</td>
<td>Summarization training;</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Participants</td>
<td>Training Intervention</td>
<td>Findings/Implications</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Mastropieri (1992)</td>
<td>Learning disabled; N= 45</td>
<td>Summarization training with a self-monitoring component; Traditional instruction.</td>
<td>Learning disabilities trained in summarization procedures performed statistically higher on all dependent measures. On one far-transfer measure, students who were trained in the monitoring component statistically outperformed those with only the summarization training.</td>
<td></td>
</tr>
<tr>
<td>Jitendra &amp; others (1998)</td>
<td>Grade 6; N=4.</td>
<td>Experimental group (N=3): received a direction main idea summarization program. Control subject (N=1): no training.</td>
<td>Results indicated the main idea instructional program produced increases in identifying and generating main ideas, with even higher levels of performance following self monitoring instruction. Student interviews confirmed the practical usefulness of the strategy in enhancing understanding textual information.</td>
<td></td>
</tr>
<tr>
<td>Klingner, Vaughn &amp; Schumm (1998)</td>
<td>Grade 4 N= 141</td>
<td>Control condition: received researcher-led instruction; Experimental condition: taught by the researchers to apply reading comprehension strategies (&quot;preview,&quot; &quot;click and clunk,&quot; &quot;get the gist,&quot; and &quot;wrap up&quot;).</td>
<td>Outcome measures, including a standardized reading test, social studies unit test, audiotapes of group work, indicated that students in the experimental condition made greater gains in reading comprehension and equal gains in content knowledge. Students implemented the clarification (click and clunk) and main idea (get the gist) strategies the most consistently and effectively.</td>
<td></td>
</tr>
<tr>
<td>Guthrie, Anderson &amp; Rinehart (1999)</td>
<td>Grade 5 N=53; Grade 3 N=67</td>
<td>Concept-Oriented Reading Instruction (CORI) group; Traditional instruction group.</td>
<td>The CORI context increased strategy use, conceptual learning, and text comprehension more than traditional instruction, when background was controlled.</td>
<td></td>
</tr>
<tr>
<td>Cordero-Ponce (2000)</td>
<td>Undergraduate; N=15</td>
<td>Meta-cognitive strategy training in summarizing on the ability of foreign learners</td>
<td>Students made substantial progress after training significantly more ideas in their recall improved their ability to use summarization rules maintain 3-wk follow-up</td>
<td></td>
</tr>
<tr>
<td>Jitendra, Hoppes, Xin (2000)</td>
<td>Middle school learning &amp; behavioral disabled; N=33</td>
<td>Experiment group: generate main ideas statements using main idea strategy instruction and a self-monitoring procedure.</td>
<td>On the training measure, instructional procedure increased reading comprehension of experimental group. On near transfer measure, experimental group outperformed posttest and delay posttest items requiring selection responses. 6 wks later, students maintained strategy usage on selection type responses on the near transfer measure but not far transfer measure.</td>
<td></td>
</tr>
<tr>
<td>Ferguson (2001)</td>
<td>Grade 6 N=41</td>
<td>Summarizing strategy alone meta-cognitive strategy instruction including the value, purpose, &amp; self-monitoring of the summarizing</td>
<td>These findings suggest that meta-cognitive strategy instruction including the value, purpose, &amp; self-monitoring of the summarizing strategy is more effective in increasing reading comprehension than the summarizing strategy alone.</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Participants</td>
<td>Design</td>
<td>Strategy</td>
<td>Findings</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------------------</td>
<td>-------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Wang (2001)</td>
<td>Undergraduate; N=66</td>
<td>Lecture</td>
<td>Structure questions group; summarizing group.</td>
<td>No significant difference between two strategies on transfer task</td>
</tr>
<tr>
<td>Elosua et al. (2002)</td>
<td>12 years old v.s. 16 years old</td>
<td>Independent variables: active text-processing strategies (main-idea identification and summarization) at two development levels (12 and 16 yr-olds) Dependent variables: reading span, reading time, construction of macrostructure</td>
<td>Reduction in developmental differences in the experimental groups at posttest was also expected Results showed a significant improvement in the experimental groups' reading comprehension &amp; recall.</td>
<td></td>
</tr>
<tr>
<td>Glover (2002)</td>
<td>At risk middle – school students</td>
<td>Summarization, self-questioning, mnemonic imagery, 2 or 3 strategies combination</td>
<td>No significant differences on identifying main ideas, generating inferences, and remembering information from text.</td>
<td></td>
</tr>
<tr>
<td>Hutton (2002)</td>
<td>English learners after school; N=9</td>
<td>Interview, observations, assessments, work samples, and journal entries of 9 students</td>
<td>Comprehension scores increased. Students were more positive about reading. Increased self-efficacy in regard to their skill as a reader. Summarizing and clarifying were perceived as the most useful strategies but actual use were not the same.</td>
<td></td>
</tr>
<tr>
<td>Mastropieri et al. (2003)</td>
<td>(mean age 15.7 yrs) with mild disabilities; N=16</td>
<td>World history</td>
<td>Peer-tutor instructional condition guided-notes instructional condition</td>
<td>Students in peer-tutoring condition significantly outperformed those in the guided-notes condition on content-area test. Students in the tutoring condition performed significantly better at using reading comprehension summarization strategy independently, and at remembering the strategy steps.</td>
</tr>
<tr>
<td>Philbrick (2003)</td>
<td>Grade 5 ; N=131</td>
<td>Expository text</td>
<td>Group1: meta-cognitive strategy + social studies content instruction; group 2: 4 strategies only (predicting, questioning, thinking aloud, and summarizing); group 3: social studies content only</td>
<td>Instruction in meta-cognitive strategy use resulted in significant gains in reading comprehension scores and increased meta-cognitive awareness in both the combined and the strategies only groups. The combined strategy group scored significantly higher on comprehension than either of the other groups.</td>
</tr>
<tr>
<td>Weedman (2003)</td>
<td>Grade 9</td>
<td>Class materials</td>
<td>Experimental : reciprocal teaching: only generating questions only summarizing 4 strategies: generating questions, (predicting, questioning, thinking aloud, and summarizing);</td>
<td>No significant was found from all data. 4 strategies group perform better than only one strategy group but not significantly.</td>
</tr>
<tr>
<td>Study</td>
<td>Participants</td>
<td>Context</td>
<td>Instruction/Methodology</td>
<td>Findings</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------</td>
<td>---------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Guthrie et al., (2004)</td>
<td>Grade 3; 19 classrooms</td>
<td>Reading/language arts and science</td>
<td>CORI (Concept-Oriented Reading Instruction); SI (Strategy Instruction); TI (traditional instruction).</td>
<td>Students in CORI classrooms were higher than SI and/or TI students on measures of reading comprehension, reading motivation, and reading strategies.</td>
</tr>
<tr>
<td>Hogewood (2004)^D</td>
<td>Grade 9 poor readers</td>
<td>Social studies text book</td>
<td>Traditional group by Palincsar &amp; Brown's (1984) reciprocal teaching simple version of reciprocal teaching four strategies group: predicting, clarifying, questioning, and summarizing two strategies only: questioning and summarizing.</td>
<td>Each groups improved reading comprehension scores on the measure of this study. Reciprocal teaching procedures can be adapted to make them easier to implement and the reading comprehension skills will still improved.</td>
</tr>
<tr>
<td>Lubliner (2002)^D</td>
<td>Grade 9</td>
<td>Social studies textbooks; areas most difficult for children to read</td>
<td>Control group; clarifying group; questioning group; combined questioning &amp; clarifying group.</td>
<td>The clarifying group made significant gains in reading comprehension &amp; vocabulary, compared to the control group. The combined questioning &amp; clarifying treatment had no effect on two of the experimental classes. Second language learners, including limited English proficient students &amp; children with learning disabilities, were included in the analysis &amp; made significant gains in reading comprehension compared to control group students. Self-reported clarifying strategy use was significantly correlated with reading comprehension scores.</td>
</tr>
<tr>
<td>Hess (2005)^D</td>
<td>Grade 4 and 5</td>
<td>Expository text</td>
<td>Instruction: 2 reading strategies-- clarification and summarization Teacher interviews and classroom observations</td>
<td>Students improved during 10 weeks study in the quality of their discussion of expository text. Students used more questions at a higher critical level of thinking, based on Bloom’s taxonomy. Students achieved higher comprehension test scores.</td>
</tr>
</tbody>
</table>

^D: from Dissertation Abstracts International Database
Appendix D
Recruitment Statement

Hello, my name is Ya-Ling Lan. I would like to share with you an opportunity to
volunteer to participate in an on-line research study. I am here to recruit participants for a study in
how college learners read and process text. The study will require about an hour of your time.
You will receive extra credit points (3 points) as directed by your instructor.

To participate, you will have to log on the displayed website. Once there, you will be
asked a few demographic questions. This includes entering the last digits of your Penn State ID.
This is important as it will be used to give you credit. Once you fill out the demographic
information, you will read some instructional materials about standardized testing, do a puzzle
task, and take a test.

You may log on the website from any computer starting today. You may participate any
time, day or night, up until Monday at midnight. This study does not require you to sign up for a
time or meet the researchers. You participate on your own time. If you have any questions you
may contact me through yxl222@psu.edu.
Appendix E

Demographic Questionnaire

Before we start the activity, please complete the following questions:

<table>
<thead>
<tr>
<th>Questions:</th>
<th>Your Answers:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) What are the last five digits of your Penn State ID? (e.g. 12136)</td>
<td>1) ____________</td>
</tr>
<tr>
<td>2) In what month and year were you born? (e.g. 05/1986)</td>
<td>2) ____________</td>
</tr>
<tr>
<td>3) For how many semesters have you attended university classes?</td>
<td>3) ____________</td>
</tr>
<tr>
<td>4) What is your gender? (M/F)</td>
<td>4) ____________</td>
</tr>
<tr>
<td>5) What is your current college major?</td>
<td>5) ____________</td>
</tr>
<tr>
<td>6) How many Statistics courses did you take in high school?</td>
<td>6) ____________</td>
</tr>
<tr>
<td>7) How many Statistics courses did you take at university?</td>
<td>7) ____________</td>
</tr>
<tr>
<td>8) What is your MATH SAT score?</td>
<td>8) ____________</td>
</tr>
<tr>
<td>9) Have you completed chapter 16 in EDPSY14? (Y/N)</td>
<td>9) ____________</td>
</tr>
</tbody>
</table>
Appendix F

The Diagram of Standardized Test Scores

Comparison of the various grading methods in a normal, bell-shaped curve. Includes:

- Standard deviations
- Cumulative percentages
- Percentiles
- Z-scores
- T-scores
- Standard nine (Stanines)
- Percentage in stanine
Appendix G

The Text of Standardized Test Scores

◆ **Standardized test scores**

A *standardized test* is a test, an assessment, or an evaluation instrument that a large sample of students take under uniform conditions and are scored according to uniform procedures. Some typical examples of these tests are the Scholastic Aptitude Test (SAT), and the California Achievement Test (CAT). Most students have taken these types of tests.

◆ **Central tendency**

To interpret *standardized test scores*, basic statistics and knowledge of *central tendency* and *normal distribution* are necessary. The three most commonly used central tendency measures are the *mode*, *mean*, and *median*. The *mode* is defined as the most frequently occurring in the distribution. A distribution may have two or more modes. The *median* is the score which marks the middle point. Therefore, half of the students have scores below this point and the other half have scores above it. Most of us have learned from high school that the *mean* is the average. We also learned that the average is obtained by adding all the scores in the group and dividing the sum by the number of students in the group. Among the measures of central tendency, the mean is the most commonly reported.

◆ **Standard deviation**

In addition to central tendency, the *variance* or *standard deviation* is calculated for every test. The variance is calculated through a series of steps. Given a frequency distribution, we first calculate the *mean*. Second, the mean is subtracted from each of the individual scores. The result is then squared to eliminate negative numbers. These numbers are then totaled. Next, the sum is divided by the total number of students and the result is the value of *variance*. The *standard deviation* is the square root of the *variance*. Usually, the standard deviation is denoted by $\sigma$ and the variance is denoted by $\sigma^2$. The purpose of standard deviation is to provide a measure of how spread out the scores are in a given distribution. The mean and standard deviation together provide a simple and efficient description of a distribution.

◆ **Normal distribution**

Because standardized tests are administered to large numbers of students, the resulting distribution of scores approximates a *normal distribution*, also known as a *normal curve*, or a *bell-shaped curve*. The characteristics of a normal distribution include a symmetric distribution with the mean, median, and mode as the same value in the center. If these measures of central tendency are not the same, the resulting distribution is considered asymmetric, or skewed.

◆ **Raw scores**

A *raw score* is the most basic score. A raw score simply represents the number of items answered correctly by a student. For example, a student who received a 39 on a 52-item test, receives a raw score of 39. Because raw scores do not provide the ability to compare across tests, measures of relative position, such as *percentiles*, *stanines*, and *standard scores* such as *z-scores* or *T-scores* are more often used when interpreting standardized tests.
Percentiles

Percentiles (also called percentile ranks or percentile scores) are also reported for standardized tests. A student’s percentile indicates the % of students whose raw scores are below that student in the norm group.

z-score

The most convenient standardized testing scoring system is the z-score. The z-score is the number of standard deviations a student’s raw score is from the mean of the norm group. That is, the z-score of a student is the student’s deviation score expressed in units of standard deviation. If the z-score is positive, it indicates a score above the mean; if the z-score is negative, it indicates a score below the mean. Note that the mean of a z-score is always 0 and the standard deviation of a z-score is always 1. The possible z-scores range between -3 and 3. For example, given an exam for a class in which the mean score is 40 and the standard deviation is 5. If Alice’s raw score is 44, Alice’s z-score is (44 – 40) / 5 = 0.8. After all students’ raw scores for this class are converted to z-scores, the mean z-score will be 0 and standard deviation of the distribution will be 1.

When the raw scores are normally distributed, we can convert z-scores to percentiles scores. That is, given a z-score, we can find what proportion of the students in this group fall below this z-score. For example, a student with a z-score of 0 would be in the 50th percentile; a student with a z-score of 1 would be in the 84th (= 50th + 34th ) percentile; a student with a z-score of 2 would be in the 98th percentile; while a student with a z-score of -1 would be in the 16th (= 50th - 34th ) percentile; a student with a z-score of -2 would be in the 2nd (= 50th - 34th -14th ) percentile. Therefore, Alice’s z-score of 0.8 would place her around the 79th percentile.

T-score

Another popular standardized testing score is the T-score. The T-score is similar to the z-score. The T-score has a mean of 50, however, instead of 0 as with the z-score, and a standard deviation of 10 instead of 1 as with the z-score. If a T-score is over 50, it indicates the score is above the mean; if a T-score is under 50, it indicates a score is below the mean. For example, Alice scores 0.8 standard deviations above the mean on a test; thus, her z-score is 0.8. After transforming her z-score to a T-score, she gets a T-score of (10*0.8) + 50 = 58.

Any given score can be interpreted as either a z-score or a T-score. For example, a T-score of 50 equals a z-score of 0; a T-score of 60 equals a z-score of 1; a T-score of 40 equals a z-score of -1; and so forth.

Stanine

While percentile ranks provide a single score, another type of standardized test score that is often reported is the stanine, or standard nine. This indicates a range of scores. There are nine stanines. Stanine 5 is in the middle of the normal distribution and includes all scores within one-fourth of a standard deviation both above and below the mean. Representing scores below the mean, stanines 4, 3, and 2 are each one-half of a standard deviation in width. Similarly, stanines 6, 7, and 8 are one-half a standard deviation in width above the mean. For example, stanine 5 represents a z-score between -.25 and .25; stanine 6 represents a z-score between +.25 and +.75; and so forth. Stanines 1 and 9 cover the rest of the lower tail and upper tail respectively. Stanine scores also used to provide a broad measure of test performance.
Appendix H

Free Recall

You just studied a passage and a diagram about standardized test scores. Please write down every thing you can remember from both the passage and the diagram.

<< You can use the back of the page. >>
<< If you finish, please wait for the investigator to give you the next instruction >>
Appendix I

Multiple-Choice Recognition Test

You just studied a passage and a diagram about standardized test scores. For each question, please choose the response you believe provides the best answer and write your response in the box (A - D).

Your answer:

1. A student scored at the 88th percentile in mathematics achievement on the national assessment. Of the following, which is an appropriate interpretation of the student's score?
   A. His T-score is approximately 48.
   B. His z-score is less than 1.
   C. His stanine score is approximately 6.
   D. His score is higher than the mean.

2. Kevin received a raw score of 56. Given a mean of 50, a median of 50, a standard deviation of 4, what is Kevin's z-score?
   A. -1.5
   B. 0.0
   C. 1.5
   D. 2.5

3. Which of the following is most closely represented by a z-score?
   A. Mode
   B. Mean
   C. Percentile
   D. Standard deviation

4. A score at the mean of the normal distribution represents a stanine of
   A. 0
   B. 1
   C. 5
   D. 9

5. Given a person who has a T score of 30 on a norm-referenced assessment, what can we conclude regarding his performance compared to the norm group? The student:
   A. did well on the assignment.
   B. performed average on the assignment.
   C. performed within 1 standard deviation of the mean.
   D. performed poorly on the assignment.

<< Items continue on the following page. ➔➔➔ >>
6. In a normal distribution with a mean of 68 and a standard deviation of 4, 84% of scores fall below a score of approximately
   A. 64.
   B. 72.
   C. 76.
   D. 80.

7. A z-score of 2 is at which percentile rank?
   A. 34th
   B. 50th
   C. 84th
   D. 98th

Use the following information for items 8-11.
A subtest of a standardized test has a mean of 40 with a standard deviation of 4. (Assume a very large sample so the test results nearly fit a normal distribution.) Cindy scores a 46 on the subtest. Cory scores a 44 on the same subtest.

8. Based on this information, the best approximation of the following of Cindy’s percentile is:
   A. 40.
   B. 44.
   C. 84.
   D. 94.

9. Of the following the best description of Cory’s stanine is:
   A. 3.
   B. 5.
   C. 7.
   D. 9.

10. Cory’s z-score on the subtest would be:
    A. -1.0
    B. 0
    C. 1.0
    D. 1.5

11. Cory’s T-score on the subtest would be:
    A. 40.
    B. 44.
    C. 50.
    D. 60.

<< Items continue on the following page. >>
12. If Lisa scored a 25 on a standardized test with a mean of 20 and a standard deviation of 5, about what percent of the students who took the test had scores lower than hers?
   A. 95%
   B. 85%
   C. 75%
   D. 65%

13. If Nick got a raw score of 26 on a standardized test with a mean of 28 and a standard deviation of 8, about what percent of the students taking the test scored lower than he did?
   A. 10%
   B. 20%
   C. 40%
   D. 55%

14. In a distribution with a mean of 50 and a standard deviation of 5, a T-score of 30 equals a raw score of
   A. 35.
   B. 40.
   C. 45.
   D. 55.

15. Given a person who has a T-score of 70 on a standardized test, what can we conclude regarding his performance compared to the norm group? The student:
   A. performed significantly above average on the assessment.
   B. performed at about an average level on the assessment.
   C. performed within 1 standard deviation of the mean.
   D. performed significantly below average on the assessment.

16. On a recent standardized test, Becky had a z-score of 2. What is the closest assessment of her performance compared to the group?
   A. Her T-score is negative.
   B. Her percentile rank is below 85.
   C. Her percentile rank is above 85.
   D. Her stanine score is 4.

17. Jordan’s score on a standardized test of mathematical aptitude was at the 85th percentile. Which of the following best describes this score?
   A. Jordan got 85% of the test items right.
   B. Jordan’s rank in the class is about 15th.
   C. 15% of all students taking this test scored higher than Jordan.
   D. For any given test item, Jordan’s chances of answering it correctly are 8.5 out of 10.
18. In a normal distribution, a T-score of 50 is at which percentile?
A. 84th  
B. 50th  
C. 34th  
D. 16th

19. You have two sets of scores, which are as follows:
Set A: Mean—40 Median—40 standard deviation—4
Set B: Mean—40 Median—41 standard deviation—6
Of the following, which is the most accurate statement?
A. The students’ performance for both sets was equal since the means are the same.
B. The scores in Set B tend to be spread out more than the scores in Set A.
C. The students in Set B actually performed better since the standard deviation is higher than in Set A.
D. The students in Set B actually performed better since both the median and standard deviation are higher than in Set A.

20. What is the cumulative percentage of z-score = -2?
A. 2.3%  
B. 15.9%  
C. 50%  
D. 84.1%
Appendix J

Activity: Notetaking

In this activity, you will study a diagram about standardized test scores. Please take notes about the diagram similar to the notes you take when you study class materials. You may write words or sentences in the response box on the right side of this page. You may refer to the diagram while you are taking notes.

<< You can use the back of the page. >>
<< If you finish, please wait for the investigator to give you the next instruction >>
Appendix K

Activity: Self-Questioning

In this activity, you will study a diagram about standardized test scores. Please write down questions you would ask yourself to help yourself understand the diagram as if you were studying for a class. You may generate questions in the response box of this page. You may refer to the diagram while you are writing questions.
Appendix L

Activity: Summarizing

In this activity, you will study a diagram about standardized test scores. Please summarize what you learned from the diagram as you might summarize information when you were learning in a class. You may write words or sentences in the response box of this page. You may refer to the diagram while you are summarizing.

<< You can use the back of the page. >>

<< If you finish, please wait for the investigator to give you the next instruction >>
Appendix M

Idea Units of Free Recall

<table>
<thead>
<tr>
<th>No</th>
<th>The content of ideas</th>
<th>From the text only</th>
<th>From diagram or text</th>
</tr>
</thead>
<tbody>
<tr>
<td>01.</td>
<td>A <em>standardized test</em> is a test, an assessment, or an evaluation instrument</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>02.</td>
<td>A <em>standardized test</em> is an evaluation that a large sample of students take.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>03.</td>
<td>A <em>standardized test</em> is an evaluation that students take under uniform conditions and are scored according to uniform procedures.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>04.</td>
<td>Some typical examples of these tests is the Scholastic Aptitude Test (SAT)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>05.</td>
<td>Some typical examples of these tests is the Scholastic the California Achievement Test (CAT).</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>06.</td>
<td>One of the most commonly used central tendency measures is the <em>mode</em>.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>07.</td>
<td>One of the most commonly used central tendency measures is the <em>median</em>.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>08.</td>
<td>One of the most commonly used central tendency measures is the <em>mean</em>.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>09.</td>
<td>The <em>mode</em> is defined as the most frequently occurring in the distribution.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>A distribution may have two or more modes.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>The <em>median</em> is the score which marks the middle point.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>Half of the students have scores below this point and the other half have scores above it.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td>The <em>mean</em> is the average.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td>The average is obtained by adding all the scores in the group and dividing the sum by the number of students in the group.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>15.</td>
<td>Among the measures of central tendency, the mean is the most commonly reported.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>16.</td>
<td>Given a frequency distribution, we first calculate the <em>mean</em>. Second, the mean is subtracted from each of the individual scores. The result is then squared to eliminate negative numbers. These numbers are then totaled. Next, the sum is divided by the total number of students and the result is the value of <em>variance</em>.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>17.</td>
<td>The <em>standard deviation</em> is the square root of the <em>variance</em>.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>18.</td>
<td>The standard deviation is denoted by $\sigma$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>19.</td>
<td>The variance is denoted by $\sigma^2$.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>20.</td>
<td>The purpose of standard deviation is to provide a measure of how spread out the scores in a given distribution.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>21.</td>
<td>The mean and standard deviation together provide a simple and efficient description of a distribution.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>22.</td>
<td>The resulting distribution of <em>standardized test</em> scores approximates a <em>normal distribution</em>.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>23.</td>
<td>A <em>normal distribution</em> is also known as a <em>normal curve</em>.</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
24. A **normal distribution** is also known as a **bell-shaped curve**.

25. The characteristics of a normal distribution include a symmetric distribution.

26. A normal distribution is a distribution with the mean, median, and mode as the same value in the center.

27. If these measures of central tendency are not the same, the resulting distribution is considered asymmetric, or skewed.

28. A **raw score** is the most basic score.

29. A raw score simply represents the number of items answered correctly by a student.

30. For example, a student who received a 39 on a 52-item test, receives a raw score of 39.

31. Raw scores do not provide the ability to compare across tests.

32. Measures of relative position such as **percentiles** are more often used when interpreting standardized tests.

33. Measures of relative position such as **stanines** are more often used when interpreting standardized tests.

34. Measures of relative position such as **standard scores** are more often used when interpreting standardized tests.

35. Measures of relative position such as **z-scores** are more often used when interpreting standardized tests.

36. Measures of relative position such as **T-scores** are more often used when interpreting standardized tests.

37. Percentiles are also called **percentile ranks**.

38. Percentiles are also called **percentile scores**.

39. A student’s **percentile** (or cumulative percentage) indicates the % of students whose raw scores are below that student in the norm group.

40. The most convenient standardized testing scoring system is the **z-score**.

41. The **z-score** is the number of standard deviations a student’s raw score is from the mean of the norm group.

42. The z-score of a student is the student’s deviation score expressed in units of standard deviation.

43. If the z-score is positive, it indicates a score above the mean.

44. If the z-score is negative, it indicates a score below the mean.

45. The mean of a z-score is always 0.

46. The standard deviation of a z-score is always 1.

47. The possible z-scores range between -3 and 3.

48. For example, given an exam for a class in which the mean score is 40 and the standard deviation is 5. If Alice’s raw score is 44, Alice’s z-score is \((44 – 40) / 5 = 0.8\).

49. After all students’ raw scores are converted to z-scores, the mean z-score will be 0.

50. After all students’ raw scores are converted to z-scores, the standard deviation of the distribution will be 1.

51. When the raw scores are normally distributed, we can convert z-scores to **percentiles scores**.

52. Given a z-score, we can find what proportion of the students in this group fall below this z-score.

53. A student with a z-score of 0 would be in the 50th percentile.
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>54.</td>
<td>A student with a z-score of 1 would be in the 84&lt;sup&gt;th&lt;/sup&gt; (= 50&lt;sup&gt;th&lt;/sup&gt; + 34&lt;sup&gt;th&lt;/sup&gt;) percentile.</td>
</tr>
<tr>
<td>55.</td>
<td>A student with a z-score of 2 would be in the 98&lt;sup&gt;th&lt;/sup&gt; percentile.</td>
</tr>
<tr>
<td>56.</td>
<td>A student with a z-score of -1 would be in the 16&lt;sup&gt;th&lt;/sup&gt; (= 50&lt;sup&gt;th&lt;/sup&gt; - 34&lt;sup&gt;th&lt;/sup&gt;) percentile.</td>
</tr>
<tr>
<td>57.</td>
<td>A student with a z-score of -2 would be in the 2&lt;sup&gt;nd&lt;/sup&gt; (= 50&lt;sup&gt;th&lt;/sup&gt; - 34&lt;sup&gt;th&lt;/sup&gt; - 14&lt;sup&gt;th&lt;/sup&gt;) percentile.</td>
</tr>
<tr>
<td>58.</td>
<td>Alice’s z-score of 0.8 would place her around the 79&lt;sup&gt;th&lt;/sup&gt; percentile.</td>
</tr>
<tr>
<td>59.</td>
<td>The T-score is similar to the Z-score.</td>
</tr>
<tr>
<td>60.</td>
<td>The T-score has a mean of 50, (however, instead of 0 as with the z-score.)</td>
</tr>
<tr>
<td>61.</td>
<td>The T-score has a standard deviation of 10 (instead of 1 as with the z-score.)</td>
</tr>
<tr>
<td>62.</td>
<td>If a T-score is over 50, it indicates the score is above the mean.</td>
</tr>
<tr>
<td>63.</td>
<td>If a T-score is under 50, it indicates a score is below the mean.</td>
</tr>
<tr>
<td>64.</td>
<td>For example, Alice scores 0.8 standard deviations above the mean on a test;</td>
</tr>
<tr>
<td>65.</td>
<td>After transforming Alice’s z-score to a T-score, she gets a T-score of (10*0.8) + 50 = 58.</td>
</tr>
<tr>
<td>66.</td>
<td>Any given score can be interpreted as either a z-score or a T-score.</td>
</tr>
<tr>
<td>67.</td>
<td>A T-score of 50 equals a z-score of 0;</td>
</tr>
<tr>
<td>68.</td>
<td>A T-score of 60 equals a z-score of 1;</td>
</tr>
<tr>
<td>69.</td>
<td>A T-score of 40 equals a z-score of -1; (and so forth).</td>
</tr>
<tr>
<td>70.</td>
<td>The stanine is also called standard nine.</td>
</tr>
<tr>
<td>71.</td>
<td>The stanine or standard nine indicates a range of scores.</td>
</tr>
<tr>
<td>72.</td>
<td>There are nine stanines.</td>
</tr>
<tr>
<td>73.</td>
<td>Stanine 5 is in the middle of the normal distribution and includes all scores within one-fourth of a standard deviation both above and below the mean.</td>
</tr>
<tr>
<td>74.</td>
<td>The interval between each stanine score are one-half of a standard deviation in width.</td>
</tr>
<tr>
<td>75.</td>
<td>Representing scores below the mean, stanines 4, 3, and 2 are each one-half of a standard deviation in width.</td>
</tr>
<tr>
<td>76.</td>
<td>Stanines 6, 7, and 8 are one-half a standard deviation in width above the mean.</td>
</tr>
<tr>
<td>77.</td>
<td>Stanine 5 represents a z-score between -.25 and +.25.</td>
</tr>
<tr>
<td>78.</td>
<td>Stanine 6 represents a z-score between +.25 and +.75; (and so forth).</td>
</tr>
<tr>
<td>79.</td>
<td>Stanines 1 and 9 cover the rest of the lower tail and upper tail respectively.</td>
</tr>
</tbody>
</table>

**total** | **41** | **38**