

**THE EFFECTS OF IMMEDIATE AND DELAYED
SUMMARIZING ON JUDGMENTS OF LEARNING,
STUDYING PATTERNS AND ACHIEVEMENT**

by

Wenting Ma
MSc. Simon Fraser University, 2006

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

In the
Faculty
of
Education

© Wenting Ma 2008

SIMON FRASER UNIVERSITY

Fall 2008

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without permission of the author.

APPROVAL

Name: Wenting Ma
Degree: Master of Arts
Title of Thesis: The Effects of Immediate and Delayed Summarizing on Judgments of Learning, Studying Patterns and Achievement

Examining Committee:

Chair: Jeff Sugarman, Associate Professor,
Faculty of Education

Philip Winne, Professor, Faculty of Education
Senior Supervisor

John Nesbit, Associate Professor, Faculty of Education
Committee Member

Kevin O'Neill, Associate Professor, Faculty of Education, SFU
External Examiner

Date Defended/Approved:

December 8, 2008

Declaration of Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website <www.lib.sfu.ca> at: <<http://ir.lib.sfu.ca/handle/1892/112>>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada



SIMON FRASER UNIVERSITY
THINKING OF THE WORLD

STATEMENT OF ETHICS APPROVAL

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

(a) Human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

(b) Advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

(c) as a co-investigator, in a research project approved in advance,

or

(d) as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Bennett Library
Simon Fraser University
Burnaby, BC, Canada

ABSTRACT

This study explored whether writing a summary immediately or with delay after studying can improve learners' metacomprehension accuracy, influence their regulation of study, and enhance reading comprehension immediately and under transfer conditions. Undergraduates (N=75) were randomly assigned to three groups: no summary, immediate summary and delayed summary. After screening data to remove participants who did not apply plausible effort, no statistically detectable differences were found for test performance across groups nor was transfer of the hypothesized effect observed. I conclude there is need to consider difficulties of reading texts and tests, and learners' metacomprehension skills to achieve a high level of metacomprehension accuracy and affect studying productively.

Keywords: metacomprehension, self-regulated learning, judgment of learning, restudy

DEDICATION

To my parents and Lei

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my profound thanks to all the people who have helped make this research possible. First, I would like to express my sincere gratitude and thanks to my senior supervisor Dr. Philip H. Winne for his consistent patience and guidance to support me through this period. Thank you for always pushing me to do my best work. I would also like to thank Dr. John C. Nesbit for being part of my supervisory committee and supportive for my work.

I would also like to thank Mrs Michele Secret who offered generous help and advice to review the reading materials and questions used in my study. In addition, I would like to thank Mr. Dave Lampron and Mr. Elliot Wang for their consistent support during the period while I was writing up my thesis.

Last, but certainly not least, I would like to express my special thanks to my parents and Lei. I would not have made all this through without their encouragement and support. Thank you for being always so patient and so supportive with me.

TABLE OF CONTENTS

Approval	ii
Abstract.....	iii
Dedication	iv
Acknowledgements	v
Table of Contents	vi
List of Figures.....	viii
List of Tables	ix
CHAPTER 1: INTRODUCTION.....	1
1.1 Purpose of This Research	1
1.2 The Research Reported in This Study	2
1.3 Thesis Structure	3
CHAPTER 2: LITERATURE REVIEW	5
2.1 Overview of Self-regulated Learning (SRL).....	5
2.1.1 The Cognitive and Social-Cognitive SRL Models	6
2.1.2 SRL in Computer-Based Environments	8
2.2 Judgment of Learning (JOL)	8
2.3 Metacomprehension Accuracy	12
2.3.1 Constraints on Metacomprehension Accuracy	13
2.4 Writing Delayed Summaries can Improve Metacomprehension Accuracy.....	15
2.4.1 Writing Summaries Improves Comprehension	15
2.4.2 Writing Summaries Improves Metacomprehension Accuracy	16
CHAPTER 3: METHODOLOGY	18
3.1 Method.....	18
3.1.1 Participants	18
3.1.2 Materials	19
3.1.3 Treatments	21
3.1.4 Procedure	24
CHAPTER 4: DATA ANALYSES AND RESULTS	26
4.1 Overview of the Types of Data Collected	26
4.2 Data Analyses and Results	26
4.2.1 The Analyses of Achievement Tests	27
4.2.2 The Analyses of the Time for Activities recorded in gStudy log files and Results.....	32

4.2.3	The Analyses of Summaries for the Immediate and Delayed Groups and Results	38
4.2.4	The Analyses of Questionnaires and Results	41
CHAPTER 5: DISCUSSIONS.....		47
5.1	Results Discussions	47
5.1.1	Metacomprehension Accuracy	47
5.1.2	The Test Performances among the Three Groups	47
5.1.3	The Quality of Summaries Written in the Immediate and Delayed Groups.....	48
5.1.4	The Effects of Restudy on Achievement	49
5.1.5	Attitudes toward Using Summaries and Restudy as Strategy in the Current Study and in the Future.....	50
5.2	Conjectures for Results and Possible Explanations.....	51
5.2.1	Did Participants Spend Sufficient Time Studying and Taking Tests?	51
5.2.2	Were the Reading Materials Too Difficult for Participants to Understand?	52
5.2.3	Were the Tests Too Difficult?	53
5.2.4	Metacomprehension Accuracy Needs Practice	53
CHAPTER 6: CONCLUSION AND FUTURE WORK		56
6.1	Conclusion.....	56
6.2	Limitation and Future Research Work	57
Reference List.....		59
Appendices.....		66
Appendix 1		67
Appendix 2		75
Appendix 3		80
Appendix 4		82

LIST OF FIGURES

Figure 3.1	A screenshot of the cognitive tool gStudy	21
Figure 3.2	A labeller tool in gStudy	22
Figure 3.3	The log file recorded during reading session in gStudy.....	22

LIST OF TABLES

Table 4.1	Reliability Test on Three Tests	27
Table 4.2	Descriptive Statistics of the Mean of the Two Tests in the Three Groups	28
Table 4.3	Test of Homogeneity of Variances.....	29
Table 4.4	Descriptive One-way ANOVA of the Mean of Two Tests in Three Groups	29
Table 4.5	The Descriptive Statistics of the Mean of the Three Ratings in the Three Groups.....	30
Table 4.6	Test of Homogeneity of Variances of Three Ratings.....	30
Table 4.7	One-way ANOVA of the Mean of the Three Ratings in Three Groups	31
Table 4.8	Descriptive Statistics of Time of Activities in Three Groups	33
Table 4.9	Test of Homogeneity of Variances on Activity Variables	35
Table 4.10	One-way ANOVA test for the Mean Time of Activities Spent in Three Groups.....	35
Table 4.11	Multiple Comparisons Test for Time Spent on Activities in Three Groups	36
Table 4.12	Descriptive Statistics of the Mean of Time (Summaries) in Two Groups	37
Table 4.13	The t-test Result of the Mean of the Time (Summaries) in Two Groups	37
Table 4.14	Model Summaries in Reading 1, 2 and 3 in Two Groups	39
Table 4.15	The Descriptive statistics of Summary Scores in Immediate and Delayed Groups.....	39
Table 4.16	The t-test Result of the Mean for the Immediate and Delayed Groups.....	40
Table 4.17	Participants' opinions on Restudy Areas in Q5	42
Table 4.18	Participants' opinions on Restudy Areas in Q6	42
Table 4.19	Restudy and Summary Attitude in the Three Groups	43
Table 4.20	Test of Homogeneity of Variances on Q7 and Q9	44
Table 4.21	One-way ANOVA Test for Opinions of Q7 and Q10 in Groups.....	44
Table 4.22	T-test for Comparing Opinions of Q8, Q9 and Q11 Across Groups.....	45

CHAPTER 1: INTRODUCTION

1.1 Purpose of This Research

The development of web-based technologies and open-ended online learning environments has received a lot of attention as a potent tool to remove temporal and geographical constraints for learners. It is possible nowadays that courses created in one place can be delivered to a group of learners who are separated in time or across locations. Thus, learning can occur at the same time in different places or at the same place at different times or even at different times and different places (Dabbagh & Kitsantas, 2005).

However, as many researchers have acknowledged, compared to the typical classroom, teaching and scaffolding can be different in web-based environments (e.g. Ley & Young, 2001; Niemi, Nevgi, & Virtanen, 2003; Kauffman, 2004; van den Boom, Paas, van Merriënboer, & van Gog, 2004; Dabbagh & Kitsantas, 2005; Narciss, Proske and Koerndle, 2007). The nature of many web-based environments precludes face-to-face teacher-student interaction due to the physical separation of instructors and learners. Learners, thus, are expected to be more effective in regulating methods they use when studying (Kaufman, 2004).

This study explores the effects of instructing learners to write summaries to support self-regulated learning in a web-based environment. If it is possible to improve learners' metacomprehension accuracy and enhance their self-regulated learning (SRL)

process, it is expected the achievement will improve. Such a finding affords opportunities for instructional designers and course instructors to improve their designs to support individual learners' self regulation in web-based learning environments.

1.2 The Research Reported in This Study

Some research studies have indicated that writing delayed summaries can improve learners' self monitoring, or metacomprehension accuracy, concerning how well they understand learning materials (Anderson & Thiede, 2008, Thiede & Anderson, 2003, Dunlosky & Amanda, 2005). As a result, learners then can regulate their study by allocating more time and deploying appropriate strategies to restudy less-well learned content to achieve their learning goals.

In the current study, the effects of immediate and delayed summarization are explored empirically. An experiment was conducted in a computer-based learning environment to investigate this hypothesis by examining whether learners instructed to write immediate or delayed summaries can improve strategic control for future studying activities compared to learners who did not write summaries.

Anderson and Thiede (2008) indicated that the effect on learning of writing delayed summaries failed to transfer – participants did not apply their skills to other contexts. But, Anderson and Thiede did not analyze further why this transfer did not occur. As no research known to me evaluated whether metacomprehension tactics are re-applied in a future studying session, this was examined by observing learners' behaviours with the support of an event-logging software tool.

Several conjectures were then offered to explain the possible reasons leading to the failure of seeing a delayed-summary effect in the study.

1.3 Thesis Structure

Chapter 2 is a review of research literature on the theories of self-regulated learning, judgment of learning and metacomprehension. The theory of self-regulated learning (SRL) and its cognitive and social cognitive models of SRL are briefly reviewed at the beginning of the chapter as the foundation of the current research study. The impact of the development of web-based technologies on SRL is then discussed to present the need to support SRL for learners in computer-based learning. Next, theories of judgment of learning (JOL) are described and the findings of the effects on making a delayed JOL are elaborated with respect to improving metacognitive accuracy involved in the SRL process. The key concept metacomprehension accuracy is discussed in detail to demonstrate the relevant research paradigms on improving the metacomprehension accuracy in relation to the current study. The effects of writing delayed summaries are then further examined and proposed as a promising approach to improve learners' metacomprehension accuracy as a contributor to advancing their comprehension.

Chapter 3 offers a detailed description of the overall experiment design in which 88 students at a western Canadian university voluntarily participated in the study. They were randomly assigned to the three treatment conditions: no-summary, immediate-summary and delayed-summary. The experimental procedure, learning materials and instruments used in this research are described in detail.

In Chapter 4, data sets collected in the experiment are analyzed. There were finally 69 participant records examined in the study after screening. Statistical results are reported in terms of tasks participants completed and their achievement within the three treatment conditions in the study.

Chapter 5 discusses the study results. No statistically detectable findings were observed among three groups of treatment conditions of writing no summaries, immediate-summaries and delayed summaries. This result contradicts findings previous empirical studies. Conjectures are offered for possible causes that may have led to this contradiction.

Finally, limitations of the current study and possible improvements for future research are addressed.

CHAPTER 2: LITERATURE REVIEW

2.1 Overview of Self-regulated Learning (SRL)

Theories and models of self-regulated learning (SRL) in academic learning emerged in the mid-1980s (Bandura, 1986; Boekaerts, Pintrich, & Zeidner, 2000; Winne, 1996, Zimmerman, 1989, 2000; Whipp & Chiarelli, 2004). These theories were developed to address questions of how learners take initiatives to direct their learning process (Zimmerman, 1986, 2001). More specifically, SRL involves metacognitive processes, the use of cognitive and metacognitive strategies and motivational beliefs (Zimmerman & Tsikalas, 2005).

Models of SRL (e.g., Metcalfe, 2002; Nelson & Narens, 1990; Winne & Hadwin, 1998, Winne, 1995a) have described SRL as a goal-driven process. That is, a learner starts with a learning goal in mind. Then, he/she works toward the learning goal while monitoring progress, and regulates accordingly to optimize learning by applying a suite of strategies until the goal is achieved. SRL is a recursive and multifaceted process that dynamically changes as “knowledge accumulates and changes” (Winne, 1995b, p. 223). Learners can be informed of their learning progress by monitoring their engagement in relation to their learning goal (Winne, 1995a). Therefore, accurate monitoring plays a crucial role in the effectiveness of regulating learning (Thiede, & Anderson, 2003; Winne & Perry, 2000).

2.1.1 The Cognitive and Social-Cognitive SRL Models

Cognitive models of SRL arise from a base in information-processing theories (Corno & Mandinich, 1983; Winne & Hadwin, 1998; Winne, 2001; Whipp & Chiarelli, 2004). In Winne's SRL model (Winne & Hadwin, 1998; Winne, 2001), SRL is categorized in 4 phases: defining the task, setting goals and planning, enacting strategies and adapting metacognition. In SRL, a learner starts with an assigned task. After evaluating the task, the learner sets goals and plans various strategies to achieve the goals. Then, as these strategies are enacted, the learner monitors and evaluates progress and may make continuous adaptation as internal and external feedback about the task arises (Butler & Winne, 1995; Stone, 2000). The cognitive models emphasize the use of metacognitive strategies when learners are faced with complex and challenging academic tasks (e.g., Winne, 1996; Winne & Hadwin, 1998).

From a social-cognitive perspective, Zimmerman (1998, 2000) depicted a three-phase cyclical model of SRL: forethought, performance, and self-reflection. In this model, SRL is considered "highly context dependent" (Schunk, 1989, p. 1) and focuses on subjective, behavioral, social and environmental factors (Gunawardena & Zittle, 1997; Hiltz, 1997; Richardson & Swan, 2003; Rourke, Anderson, Garrison & Archer, 2001; Swan, 2002, 2003; Zimmerman, 1995; Whipp & Chiarelli, 2004).

The social-cognitive approach to modeling SRL describes how a self-regulated learner employs cognitive and metacognitive strategies with strong self-efficacy and motivational beliefs that are influenced by social and environmental factors (Whipp & Chiarelli, 2004). The forethought phase includes metacognitive processes (e.g., task assessment) and sources of self-motivation (e.g., task values) (Zimmerman & Tsikalas,

2005). In this phase, learners conduct a task analysis and set goals in relation to self-efficacy beliefs to accomplish the given task (Whipp & Chiarelli, 2004). The performance phase involves the use of metacognitive and behavioral strategies and self-observation (metacognitive monitoring and behavioral recording) (Zimmerman & Tsikalas, 2005). Learners use a range of strategies while monitoring progress toward goal achievement. In the self-reflection phase, learners use self-judgments and self-reactions to evaluate their performance. In particular, the self-reflection phase “involves not only metacognitive self-evaluation but also affective and motivational reactions to one’s self-regulatory efforts...and adaptive rather than defensive self-reactions” (Zimmerman & Tsikalas, 2005, p. 1). Such self-judgments and self-reactions can then further influence forethought and performance in a cyclical loop. A typical self-regulated learner demonstrates initiative (Schunk, 1990, Stone, 2000), persistence (Pintrich & de Groot, 1990; Schunk, 1990; Zimmerman, 1986; Stone, 2000), and effort (Pintrich & de Groot, 1990; Stone, 2000).

Comparing those two models, the cognitively-oriented model emphasizes the importance of students' perceptions of the usefulness of strategies as the determining factor in their willingness to use them whereas the social cognitive perspective focuses more on the role of self-efficacy in students' motivation to use SRL strategies (Schunk, 1989)

2.1.2 SRL in Computer-Based Environments

Recent advances in web-based technologies and the Internet have greatly expanded the boundaries of traditional face-to-face classroom learning (Dabbagh & Bannan-Ritland, 2005; Hartley & Bendixen, 2001; Dabbagh & Kitsantas, 2005). Learners now have nearly ubiquitous access to global resources, the opportunity to communicate with a wider range of people, a range of interactive experiences to share and co-construct knowledge, and the flexibility to study at their own pace (Dabbagh & Kitsantas, 2005).

However, the nature of many web-based environments constrains face-to-face teacher-student interactions. For example, learners distributed globally can't have face-to-face interactions with their instructors as they are physically separated from them. Therefore, learning in web-based environments may require more independent learning where learners may need to be highly self-regulated (Kaufman, 2004).

Even if learners are self-regulated in a particular context, it can't be assumed they transfer these SRL strategies if they move from a typical classroom setting to a web-based one. As a result, high demand for self-regulation invites building a web-based environment to support SRL that emphasizes cognitive strategy use, metacognitive processing, and motivations (McLoughlin, Winnips & Oliver, 2000; Azevedo & Hadwin, 2005; Dabbagh, 2003).

2.2 Judgment of Learning (JOL)

Learners are self-regulated agents who construct knowledge (Winne, 2005). Empirical evidence on learning with computer-based environments shows that learners

experience difficulties in domains that are conceptually demanding (Azevedo, 2005). Judgments of learning (JOLs), a form of metacognitive judgment, are a pivotal factor when learners manage their study activities. JOLs are assessments learners make about their mastery of learning materials they study (Son & Metcalfe, 2005).

According to Rawson, Dunlosky, and Thiede (2000), a JOL is made by relating various cues generated during study to performance. When learners make a judgment of their understanding, a variety of cues are examined to determine how well they understand what they have studied. A sample cue would be the amount of information learners can adequately retrieve from texts at the time of making the judgment (Benjamin, Bjork, & Schwartz, 1998; Morris, 1990) or the ease or fluency of retrieving targeted information (Rawson & Dunlosky, 2002) or “how replete the retrieved information is” (Son & Metcalfe, 2005, p. 1116). Then, they evaluate qualities of those cues that they have access to, and make inferences about how well they will do on a future test based on those cues (Matvey, Dunlosky, & Guttentag, 2001). When learners find out they have studied inadequately or do not meet their learning goals, they compensate by allocating more time to less well-learned material, until they have reached their goal. Therefore, the accuracy of JOLs can have a great impact on learners’ performance if learners modulate and regulate their approach to the learning based on JOLs.

However, research shows the accuracy of learners’ JOLs is very often low (Nelson & Dunlosky, 1991). One prevalent explanation is that JOLs are usually based on the information available in learners’ memories immediately after study. According to the principle of “monitoring-dual-memory” that Nelson and Dunlosky (1991) proposed,

JOLs are made by accessing information – answers per se as well as secondary information about completeness, accuracy and ease of retrieval – stored both in short-term memory (STM) and long-term memory (LTM). When JOLs are made immediately after study, information in STM predominates and interferes with access to information in LTM. Thus, the information primarily retrieved from STM is used as the basis for JOLs. Because information accessible in STM is transitory and unreliable, relying on that information causes the low accuracy of JOLs (Son & Metcalfe, 2005).

Along with this hypothesis, many studies have shown that the accuracy of JOLs greatly improves when judgments are made in a delayed manner rather than generated immediately after studying information (for examples, see: Carroll & Nelson, 1993; Dunlosky & Nelson, 1992; Leonesio & Nelson, 1990; Lovelace, 1984a, 1984b; Nelson & Dunlosky, 1991; Weaver, & Kelemen, 1997). In a delayed JOL paradigm, a learner makes JOLs after from a few seconds to several minutes after studying information to be learned. In this case, the accuracy of a delayed JOL, measured against actual achievement, has been observed to greatly increase (Van Overschelde, & Nelson, 2006).

Three possible explanations have been suggested for improved JOL accuracy when JOLs are delayed. First, Nelson and Dunlosky (1991) suggested is that, when a judgment is made after a short delay, relevant information is retrieved primarily from LTM because the transient information in STM has vanished after a delay. As the information stored in LTM is more accurate and reliable in predicting future performance, the accuracy of JOL is improved.

A second account for the delayed JOL effect is that the context of making a delayed JOL is more similar to the context of recalling information in a test than the

context of studying materials. The similarity of the contexts optimizes the process when the delayed judgment is made and thus improves the JOL accuracy (Son & Metcalfe, 2005).

The third possible reason for improved JOL accuracy on delayed judgments concerns the increased strength of information in the memory system instead of a focus on the metamemory system (Kimball & Metcalfe, 2003; Spellman & Bjork, 1992, Son & Metcalfe, 2005). In this case, learners are assumed to attempt to retrieve information when they make a JOL. Thus, when information receives a high JOL, it has been successfully retrieved after a delay. Increased JOL accuracy therefore results, in part, because study items with high-JOLs receive more spaced retrieval attempts than low-JOL items. However, most studies of learners' JOLs involved only recalling associative information such as cued term pairs (e.g., Shaughnessy & Zechmeister, 1992). The findings of those studies failed to generalize to learners' metacognitive accuracy in comprehension monitoring of coherent texts, often called metacomprehension accuracy (Anderson & Thiede, 2008). In texts, rather than the simple relations between independent paired associates, considerable encoding and other forms of processing are involved in studying the text and making JOLs (Weaver & Kelemen, 2003). Thus, more complex monitoring strategies are required than those engaged in the associative tasks. More studies need to be conducted to identify effective methods to improve metacomprehension accuracy when learners study text materials because those kinds of materials are prevalent in schools.

2.3 Metacomprehension Accuracy

In research on reading comprehension, the monitoring process by which learners assess their comprehension and regulate study accordingly is frequently identified as a critical factor for effective learning (Anderson & Theide, 2008; Maki, Jonas, & Kallod, 1994). If learners can appropriately judge their learning to differentiate what they have learned well versus what they have not, they can allocate more resources (e.g., time) to focus on the unlearned information when they reread the learning materials. This monitoring process concerning how well a text is understood is referred to as metacognitive accuracy in comprehension monitoring or metacomprehension. Metacomprehension underlies judgments about the comprehension of reading materials (Dunlosky & Amanda, 2005) when predicting performance in the future. When learners judge that their comprehension of the reading materials is low and have not learned material well, they initiate rereading until they reach a desired performance. Thus, the effectiveness of reread is closely associated with the metacomprehension accuracy in that metacomprehension accuracy provides guidance for learners to judge what and where to study. Therefore, the higher the metacomprehension accuracy, the more effective regulation of study can be, and the better a learner's future performance will be (Anderson & Thiede, 2008).

Two types of metacomprehension accuracy are defined and measured in the current research studies. One is the relative metacomprehension accuracy measured by the correlation between judgments and test performance. High positive correlations indicate that learners are more accurate in differentiating materials from a well-learned versus a not-well-learned state. The other type is metacomprehension accuracy is

absolute metacomprehension accuracy. Absolute metacomprehension accuracy concerns whether a learner overestimates or underestimates test performance, hence yields indicators of underconfidence or overconfidence in predicting test performance. Those 2 types of metacomprehension accuracy reflect different aspects of accuracy in predicting test performance (Dunlosky & Amanda, 2005).

2.3.1 Constraints on Metacomprehension Accuracy

Improving metacomprehension accuracy theoretically is critical to effective regulation of learning. Over the past decades, the common finding in metacomprehension research is that the metacomprehension accuracy tends to be low compared to the accuracy of monitoring associative term-pairs (Anderson & Thiede, 2008; Dunlosky & Amanda, 2005; Griffin, Wiley, & Thiede, 2007).

Several factors have been revealed as factors that cause variance in metacomprehension accuracy among learners. Those factors weigh differently in different learning contexts. Understanding those factors that constrain metacomprehension accuracy will help us to combine different assumptions underlying those factors and have implications for research designed to boost metacomprehension accuracy.

One of the factors that constrains metacomprehension accuracy, as Griffin et.al (2007) suggested, is the requirement of concurrently executing two separate tasks during study, monitoring and text-comprehension. The comprehension task requires processing at the “object” level of information in the text. The monitoring task occurs at the meta-

level that focuses on one's mental processing. Also, monitoring is regarded as a secondary process to the processing at the object-level because it needs to be "informed by the object-level" (p. 93). Thus, the success of accurate monitoring relies on whether monitoring can occur concurrently with primary text processing and is available to learners in judging their progress. In most cases, when learners read texts for the first time, due to demands of object-level processing, the monitoring process at the meta-level may not be sufficiently active during reading. Lack of sufficient concurrent monitoring thus leads to the inaccuracy judgment of the comprehension level.

Another possible explanation for constraints on metacomprehension accuracy, according to Dunlosky, Rawson and Middleton (2005), is associated with the validity of cues available to make the judgment. According to the accessibility hypothesis, the judgment is based on the total accessible information at the time the judgment is made. As learners often can not evaluate the quality of information available to them, they make the judgment based on the quantity of available information. Thus, whether learners have accurate judgment depends on the correctness of the information accessed.

A third factor identified by Griffin et.al (2007) as influencing metacomprehension accuracy is individual differences. Learners vary in their comprehension abilities, such as the ease of processing texts at the object-level; and they vary in their metacomprehension abilities, such as the ability to shift from the object level to the meta-level.

As well, the nature of reading materials is also a plausible factor accounting for variance of the metacomprehension accuracy among learners. When learners cope with texts that are beyond their capabilities, even with higher metacomprehension accuracy, it

is still not viable for them to achieve high performance in tests because they simply cannot understand the texts.

2.4 Writing Delayed Summaries can Improve Metacomprehension Accuracy

2.4.1 Writing Summaries Improves Comprehension

Writing summaries has been identified as an effective method to improve learners' reading comprehension (Anderson & Thiede, 2003; Armbruster, Anderson, & Ostertag, 1987; Bean & Steenwyk, 1984; Berkowitz, 1986; Doctorow, Wittrock, & Marks, 1978; Wittrock & Alesandrini, 1990). A school of researchers suggested that when writing summaries, learners make decisions focusing on the important elements of a text and then are able to comprehend the text at a higher level than if they just passively read it (Anderson & Armbruster, 1984; Pearson & Fielding, 1996; Kintsch & van Diik, 1978; Radmacher, & Latosi-Sawin, 1995; Winograd, 1984). Other researchers claimed that generating summaries can help learners to build relations among information in the text and link object-level information to prior knowledge (Doctorow, et al., 1978; Wittrock & Alesandrini, 1990). In the summarizing processes, learners build relations between concepts and develop schemata and then transform information into a concise form (Friend, 2001).

In some other studies, summarizing text was considered a self-testing tool, helping learners to monitor their understanding of texts and enact appropriate strategies to promote understanding (Anderson & Thiede, 2003; Brown & Day, 1983; Garner, 1982; Palinscar, 1986; Palinscar & Brown, 1984; Paris & Lindauer, 1982). Those studies

suggest that writing summaries can improve the accuracy of judging one's comprehension of texts (metacomprehension accuracy), setting the stage for more effective regulation of study.

2.4.2 Writing Summaries Improves Metacomprehension Accuracy

As discussed previously, findings of research on metacomprehension accuracy suggest that writing a summary improves reading comprehension by enhancing metacomprehension accuracy (Thiede & Anderson, 2003). Compared, for example, to using key words describing texts (Thiede, Dunlosky, Griffin, & Wiley, 2005), one possible reason that summarizing can increase metacomprehension accuracy is that it requires deeper processing and understanding of a text. Another potential explanation is that summarizing plays a role in helping learners to evaluate the quality of cues they use to judge their understanding of a text (Thiede & Anderson, 2003). Thus, when learners write summaries, they can judge how retrievable information is and, as they strive to connect information to form the summary, how accurate that information is. These cues contribute to a more accurate judgment of their comprehension of the text which, in turn, affords regulating future study to remediate less well-learned information.

Similar to the effectiveness of making delayed metacognitive judgment (Dunlosky & Nelson, 1992; Maki, 1998b; Nelson & Dunlosky, 1991), the effects on improving metacomprehension accuracy were also moderated by the timing of writing summaries. According to the activation theory (Britton & Gulgoz, 1991; Fletcher, van den Broek, & Arthur, 1996; van den Broek, Ridsen, Fletcher, & Thurlow, 1996), some information remaining in short-term memory is information that is not well-understood. After a period of time, as activation decays, propositions not well-understood vanish from

short-term memory. Thus, when learners write summaries without delay after studying a text, they may access information they actually may not understand, leading to an illusion of competence in understanding the texts. On the contrary, when learners write summaries after a delay, they can only retrieve information that was well understood and stored in long-term memory. As they monitor the quality of the summary, they can examine whether it is complete and accurate and thereby generate a better basis for achieving metacomprehension accuracy.

Writing a delayed-summary, according to the construction-integration model (Kintsch, 1998), can better involve learners in linking information and connecting concepts in texts (Thiede & Anderson, 2003, Anderson & Thiede, 2008). In a series of research studies Anderson and Thiede (2003, 2008) have conducted on metacomprehension accuracy, writing delayed summaries was an effective method for boosting learners' metacomprehension accuracy. As a result of having a higher level of metacomprehension accuracy, learners can guide their restudy more effectively.

In the current study, the effects of writing delayed summaries were re-examined to investigate whether such an effect is sustainable in improving learners' metacomprehension in post-secondary education.

CHAPTER 3: METHODOLOGY

This study investigated how learners' metacognitive judgments affect their regulation of subsequent study. Learners read three texts and were randomly assigned to write summaries: immediately after studying each of the first three texts, after all three texts had been studied, or not at all. Participants also rated their comprehension of the texts they studied. Participants who wrote summaries make their ratings after writing each summary of a text. Participants not writing summaries rated comprehension immediately after reading each text. Following these activities, participants could choose to restudy any of the three texts. Making judgments of learning after reading and summarizing (or not) was predicted to influence their selections of learning materials to restudy. It was also predicted that the treatments would have differential effects on improving the accuracy of metacomprehension judgments and final test performance on a fourth text.

3.1 Method

3.1.1 Participants

In this study, 88 student volunteers were recruited campus-wide at a western Canadian university. There were 54 females (61.3%) and 34 males (38.7%) in the sample

with ages ranging from 18 to 44 years old. Seventy-three participants (82.9%) were undergraduates and the remaining 15 (17.1%) were graduate students.

After carefully verifying the reliability and validity of the collected data, 13 participants' were removed from the pool due to incompleteness of their protocols (missing data). This led to a final sample size of 75, with 47 females (62.7%) and 28 males (37.3%).

Participants were paid \$15 for their contribution to their research.

3.1.2 Materials

The 4 texts were selected from featured articles posted on websites for *Time* magazine (<http://www.time.com>) and *The New York Times* newspaper (<http://www.nytimes.com>). The length of the first three texts used in the experiment, where participants could restudy the texts, ranged from 600 to 700 words. The 4th text was approximately 1000 words. To minimize chances that some participants might have an advantage because of particular background, the articles were chosen to present a variety of rhetorical forms and topics: a book review about a book on delivering casualty notification to the families of fallen soldiers in the United States Marine Corps, a commentary on Microsoft's bribing of customers in the search-engine market, the recent trend of shrinking food packaging in America's food industry and Coca-Cola's failure in launching its new taste product. The difficulty of the 4 articles was estimated using the Flesch-Kincaid index (Flesch, 1948). Values were: book review (11.3), Microsoft (9.3),

food packaging (12.2), and Coca-Cola (9.4), with an average Flesch-Kincaid Grade Level of 10.6. The four texts are listed in Appendix 1.

In total, three tests were created to evaluate participants' understanding of those four texts. The first two tests addressed information in the first three texts and the third test addressed the 4th reading. The items in all three tests used multiple-choice format and represented levels in Bloom's revised taxonomy of recall, low inference (comprehension) and high inference (application and analysis) (Anderson & Krathwohl, 2001).

Test 1 presented three items per article, one recall item and two high-inference items. Test 2 presented the same nine items as the 1st test plus 18 new questions for the three texts, 6 recall and 12 inference, yielding a total of 27 items. The questions and their answers of Test 2 are listed in Appendix 2.

Test 3 presented nine items based on the 4th text, three recall items and six inference items. The questions and their answers of Test 3 are listed in Appendix 3.

In addition, an 11-item questionnaire was developed to collect demographical data and participants' reflections on their re-study experience at the end of the first phase when participants read and restudied three of the texts. Specifically, participants were asked about the number of texts, areas, type of information that they restudied as well as their attitude toward the usefulness of restudying and writing summaries.

The entire questionnaire is provided in Appendix 4.

3.1.3 Treatments

Participants studied all four texts in an online learning environment named gStudy. gStudy was developed by researchers in the Learning Kit Project at Simon Fraser University. It provides learners a set of built-in tools, such as taking notes on web pages, to support learners' self-regulated study in a multimedia learning environment (See Winne et al., 2006 and Nesbit & Winne, 2007 for a full discussion of this tool.) A screenshot of gStudy is presented in Figure 3.1.

Figure 3.1 A screenshot of the cognitive tool gStudy

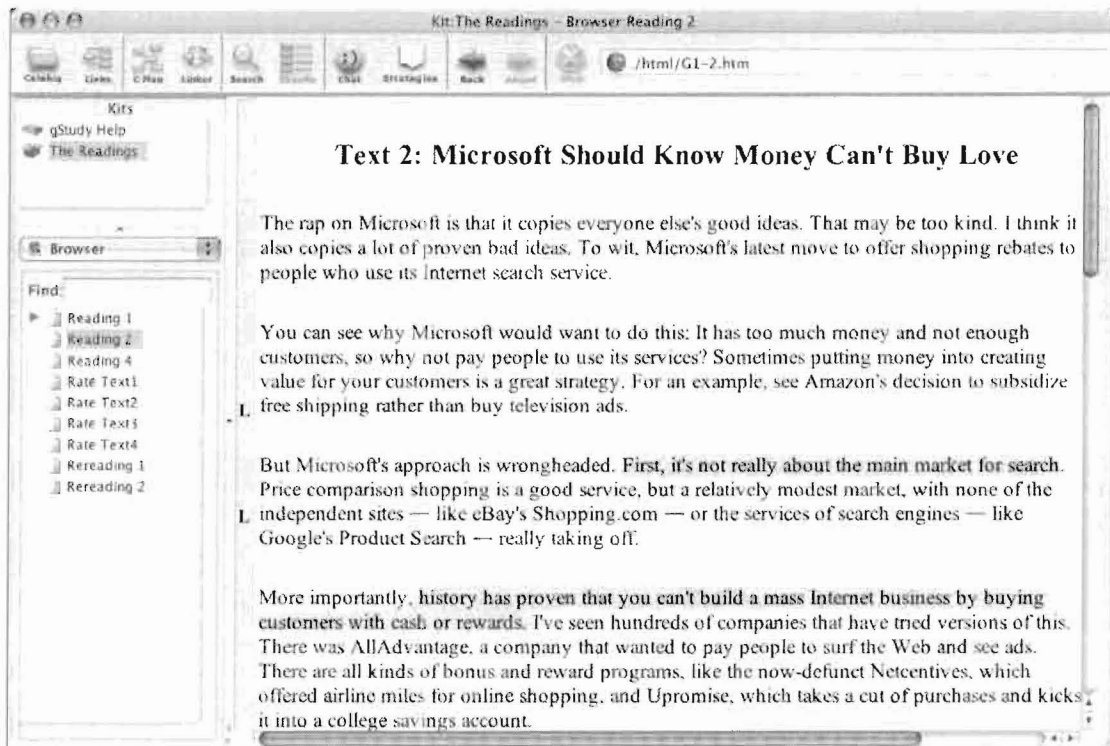


Figure 3.2 A labeller tool in gStudy

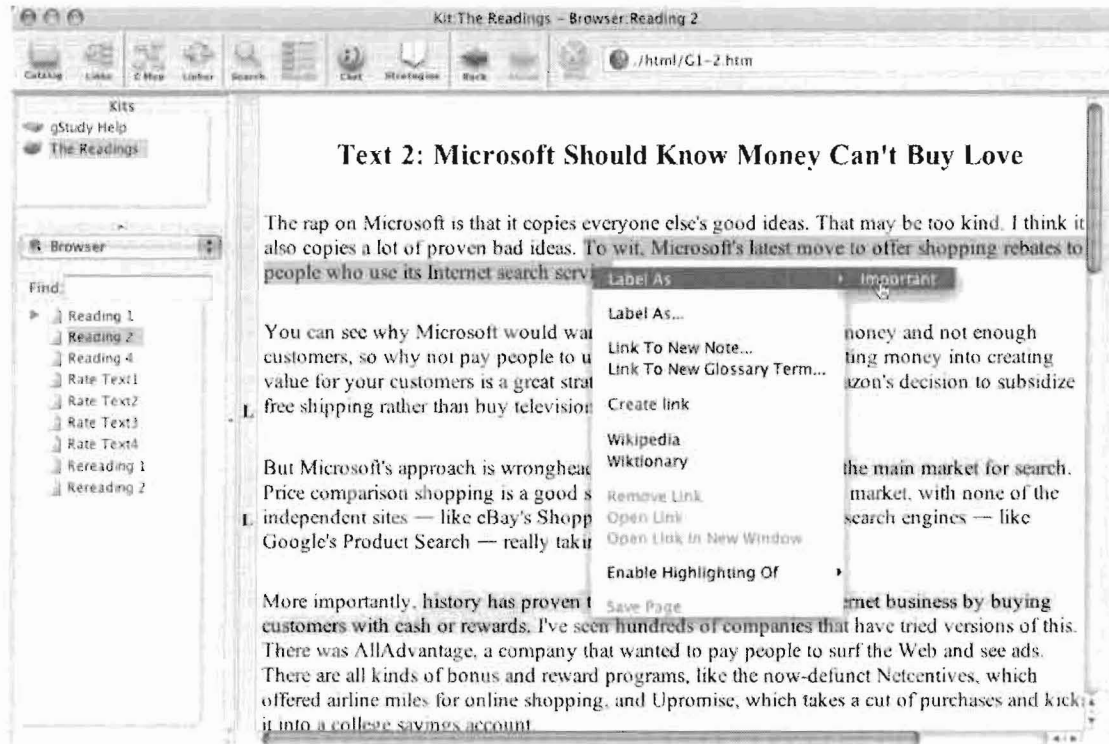


Figure 3.3 The log file recorded during reading session in gStudy

```

<dg:ScrollThumbPosition url="/html/G1-2.htm" verticalPosition="6" />
<dg:componentId ID="MainWindow.BrowserView.DetailPanel" description="Detail Panel" />
</lg:ViewEvent>
- <dg:ViewEvent container="MainWindow1" timeStamp="2008-07-31T12:09:18.062-07:00">
  <dg:ScrollThumbPosition url="/html/G1-2.htm" verticalPosition="12" />
  <dg:componentId ID="MainWindow.BrowserView.DetailPanel" description="Detail Panel" />
  </lg:ViewEvent>
- <dg:ModelEvent action="created" target="link" timeStamp="2008-07-31T12:10:10.074-07:00" user="part117">
  <dg:targetObject kitID="142585412learningkitsfuca6b5a4611b6b0332167fdc" kitName="The Readings"
  origin="395f6c4c1190da13c1f7f87">
  - <dn:link author="part117" dateCreated="2008-07-31T12:10:10.051-07:00" dateModified="2008-07-
  31T12:10:10.052-07:00" id="8707011b7a6cf2f07fe3" modifiedBy="part117" templateRef="">
    <dn:docOrigin QTmedia="false" endpointX="403" endpointY="525" image="false" rangeEnd="1749"
    rangeStart="1679" selection="the biggest motivator of behavior online has been speed and simplicity"
    startPointX="833" startPointY="506" useRange="true" />
  - <dn:destination>
    <dn:modelObject targetID="1921681101G5l0calece3610a29a0541a7fe4" moClass="toEntry"
    targetKitID="142585412learningkitsfuca6b5a4611b6b0332167fdc" />
  </dn:destination>
  </ln:link>
  - <dg:OriginMO>
  - <dc:MetaDoc target="/html/G1-2.htm" author="testqa001-a" dateCreated="2008-04-01T23:00:14.078-07:00"
  dateModified="2008-07-31T12:10:10.052-07:00" id="395f6c4c1190da13c1f7f87" modifiedBy="part117"
  name="Reading 2" templateRef="">
  - <md:links>
  - <dn:link author="part117" dateCreated="2008-07-31T12:10:10.051-07:00" dateModified="2008-07-
  31T12:10:10.052-07:00" id="8707011b7a6cf2f07fe3" modifiedBy="part117" templateRef="">
    <dn:docOrigin QTmedia="false" endpointX="403" endpointY="525" image="false" rangeEnd="1749"
    rangeStart="1679" selection="the biggest motivator of behavior online has been speed and
    simplicity" startPointX="833" startPointY="506" useRange="true" />
  - <dn:destination>
    <dn:modelObject targetID="1921681101G5l0calece3610a29a0541a7fe4" moClass="toEntry"
    targetKitID="142585412learningkitsfuca6b5a4611b6b0332167fdc" />
  </dn:destination>
  </ln:link>
  </md:links>
  
```

In this experiment, while participants studied reading texts, they were allowed to make highlights using a labeling tool in gStudy (See Figure 3.2). All participants' behaviors were traced in log files generated in gStudy (See Figure 3.3).

Participants in all groups studied the same four texts. For the first three texts, participants were required to complete their assigned tasks after completing a study session and a restudy session. In a transfer session, all participants read the 4th text and then proceeded to take the final test. All participants were informed they could take as much as time they needed and use any learning strategies freely before taking the last test.

During study and restudy sessions using the first three texts, participants were required to read each text. The order of texts was the same for all participants. In the no-summary group, participants rated their comprehension on a 0-10 scale immediately after reading each text.

In the immediate-summary group, immediately after reading the first text, participants wrote a summary and rated their comprehension of that text on a scale from 0-10. They continued this procedure until they completed all three texts.

In the delayed-summary group, participants were also required to summarize each text and rate their comprehension. However, summaries and ratings of comprehension were made after an intervening task that delayed these activities relative to prior studying. Participants read the first and then the second text, and then wrote the summary and rated comprehension for the 1st text. Then, they read the 3rd text and wrote the summary and rated comprehension for the 2nd text. After they completed the rating and summary for the 2nd text, they wrote the summary and rated comprehension for the 3rd text.

These activities took approximately 40-60 minutes depending on the particular treatment.

After taking the first test, all participants were presented with hyperlinks to the three texts for restudy. Participants were not allowed skip this restudy session and proceed directly to the 4th text. This ensured that participants would be exposed to at least one text so that data could be gathered on metacomprehension accuracy. However, there were no constraints on how they restudied any or all text(s), including whether they restudied it at all. While restudying, participants were able to access their highlights created during the first study period. They were told there was no time limit or limit on the number of times they restudied texts before heading to the 2nd test. Once participants completed restudying the first text, they were provided the links to take the 2nd test right away or continue restudying more texts. This procedure repeated until participants reviewed all texts they wanted.

When participants felt satisfied with their understanding, they proceeded to the 2nd test. Instructions for participants to write summaries were not otherwise descriptive about the product to be created. If they asked about requirements, they were told they could write a summary with a length they judge appropriate and adequate.

Participants in one group were not aware of the other treatment groups.

3.1.4 Procedure

As participants arrived at the laboratory, they randomly chose a chit of paper with a number ranging from 1 to 3. This assigned them to the no-summary group, the

immediate-summary group or the delayed-summary group. All participants signed a consent form before taking part in the study. They were informed that all data would be kept confidential and used only for research purposes. They were informed they could withdraw at any time during the study.

All participants were examined twice for the first three texts. After reading, summarizing (except for the no summary group) and rating comprehension for all three texts in a first study period, they took the 1st test. Then, they were directed to the restudy page with three hyperlinks to each of the three texts. They had to choose at least one link to one of the three texts to proceed. After they clicked up a link to a text, they went back to the reading pages they marked during their first study period. Participants were told that they could review the highlights they created before, create more highlights or just re-read the text. After they finished rereading a first text, they could either proceed to 2nd test or continue restudying the remaining texts. In all three groups, participants had full control of the number of texts to restudy. After they felt comfortable to move on, they followed the link to take the 2nd test.

After they completed the 2nd test, participants began reading the 4th text. For the 4th text, as there were no assigned tasks, all the participants proceeded to the test immediately after they completed it. After all three tests, all participants completed a questionnaire to report how they interacted with the learning materials and what they have learned in the study.

CHAPTER 4: DATA ANALYSES AND RESULTS

4.1 Overview of the Types of Data Collected

In this experiment, data collected included: test scores measuring achievement on information presented in each reading for all three groups, ratings of the three practice texts for all three groups, summaries written about practice texts by the participants of the immediate and delayed summary group, the questionnaire data for all three groups and the gStudy log file recording time stamps of activities participants performed during the experiment sessions.

Based on the previous studies Anderson and Thiede (2003, 2008) conducted to manipulate metacomprehension accuracy using delayed-summary effect, it was hypothesized the delayed-summary group would outperform the other groups on all achievement tests.

4.2 Data Analyses and Results

In this study, 75 participants generated data. However, by assuming a reading rate of 250 words/minute, which is a fast reading speed for an advanced reader (based on a casual test conducted by P. H. Winne); and 20 seconds/test item, 10 participants were judged to have unreasonably short study times or time to take tests. Looking more carefully at these 10 records, 4 actually performed above the average on all three tests. This could indicate those 4 participants took sufficient time to generate valid data. Therefore, data from those 4 participants were kept for analysis in the study. The

remaining 6 participant records were removed from the data pool, leading to a final sample size of 69 participants.

Among all 69 participants, there were 43 females (62.3%) and 26 males (37.7%). 26 participants (37.7%) were English-native speakers and the rest 43 (62.3%) participants spoke different languages such as Mandarin, Bengali and French. Fourteen (20.3%) participants were graduate students and the rest 56 (79.7%) are undergraduates, with majors varying widely from business to computing science.

Goodness-of-fit of tests were run prior to all data analyses to examine if the assumptions of normal distributions of scores were satisfied. The test results show that all data sets are approximately normally distributed.

4.2.1 The Analyses of Achievement Tests

4.2.1.1 The Reliability of the Three Tests

In Table 4.1, Cronbach's alpha was computed for the three achievement tests used in the study. The number items in the tests were 12, 27 and 9 respectively.

Table 4.1 Reliability Test on Three Tests

Test	Cronbach's Alpha	N of Items
Test 1	.21	12
Test 2	.65	27
Test 3	.58	9

The reliabilities of Test 2 and Test 3 were low but acceptable. Test 1 did not reach an acceptable level of reliability so it was removed from further analyses.

4.2.1.2 Achievement Results Across Groups

Table 4.2 shows descriptive statistics for Test 2 and Test 3. In both tests, three groups vary slightly in their mean scores on the tests. More specifically, In Test 2 and Test 3, the immediate-summary group appears to outperform the other two groups, with the no-summary group seemingly outperforming the second and the delayed-summary group the worst.

Table 4.2 Descriptive Statistics of the Mean of the Two Tests in the Three Groups

Group	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Test2 No summary	24	12.96	4.03	.82	11.26	14.66	5.00	22.00
Immediate summary	21	14.05	3.47	.76	12.47	15.63	5.00	19.00
Delayed summary	24	12.40	4.16	.85	10.66	14.17	4.00	19.00
Test3 No summary	24	5.54	2.04	.42	4.68	6.40	1.00	8.00
Immediate summary	21	6.10	1.67	.36	5.33	6.86	3.00	9.00
Delayed summary	24	5.29	1.97	.40	4.46	6.12	1.00	8.00

Table 4.3 reports the Levene test of homogeneity of variances of both tests across the three groups. Homogeneity of variances of Test 2 and Test 3 were confirmed.

Table 4.3 Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
Test2	.53	2	66	.59
Test3	.70	2	66	.50

Table 4.4 shows a one-way ANOVA for each test across groups. Results indicate there are no statistically detectable differences among means of the three groups for either Test 2 or Test 3.

Table 4.4 Descriptive One-way ANOVA of the Mean of Two Tests in Three Groups

		Sum of Squares	df	Mean Square	F	Sig.
Test2	Between Groups	30.55	2	15.27	1.00	.38
	Within Groups	1011.74	66	15.33		
Test3	Between Groups	7.48	2	3.74	1.03	.36
	Within Groups	240.73	66	3.65		

4.2.1.3 Ratings Across Groups

Table 4.5 reports descriptive statistics on ratings of comprehension made by participants in the three groups in the study. The maximum range of ratings is from 2 to 10.

Table 4.5 The Descriptive Statistics of the Mean of the Three Ratings in the Three Groups

Group		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Rating 1	No summary	24	6.96	1.90	.39	6.16	7.76	2	9
	Immediate summary	21	7.52	1.50	.39	6.84	8.21	4	10
	Delayed summary	24	7.17	1.49	.31	6.54	7.80	4	10
Rating 2	No summary	24	7.42	1.28	.26	6.88	7.96	4	9
	Immediate summary	21	7.95	1.50	.33	7.27	8.63	3	10
	Delayed summary	24	7.38	1.77	.36	6.63	8.12	4	10
Rating 3	No summary	24	7.75	1.23	.25	7.23	8.27	5	10
	Immediate summary	21	8.38	1.40	.31	7.75	9.02	5	10
	Delayed summary	24	7.79	1.96	.40	6.97	8.62	5	10

In Table 4.6, the Levene test of homogeneity of variances shows that the homogeneity of variances of Rating 2 and Rating 1 are not detectably different across the three groups. However, variances of groups' Rating 3 differ across the three groups.

Table 4.6 Test of Homogeneity of Variances of Three Ratings

	Levene Statistic	df1	df2	Sig.
Rating1	1.21	2	66	.30
Rating2	1.58	2	66	.22
Rating3	5.25	2	66	.01

In Table 4.7, one-way ANOVAs testing difference among group means on ratings indicate there were no statistically detectable differences for any of the ratings across three groups.

Table 4.7 One-way ANOVA of the Mean of the Three Ratings in Three Groups

		Sum of Squares	df	Mean Square	F	Sig.
Rating1	Between Groups	3.63	2	1.82	.67	.52
	Within Groups	179.53	66	2.72		
Rating2	Between Groups	4.55	2	2.27	.97	.38
	Within Groups	154.41	66	2.34		
Rating3	Between Groups	5.46	2	2.73	1.12	.33
	Within Groups	161.41	66	2.45		

4.2.1.4 Transfer of Effects on Metacomprehension Accuracy

In the current design, with no instruction or scaffolding provided for writing summaries, participants in all three groups were treated equivalently on the 4th reading text and, after reading this text, they completed Test 3. The transfer effects on retaining higher metacomprehension accuracy are expected to be observed for participants in the 3rd group in Test 3 as a result of a perceived improvement in their own metacomprehension accuracy by writing delayed summaries and reapplying it in the further study. Thus, if participants in Group 3 have performed better than the other two groups in Test 2, it is sensible to find out if the 3rd group would also perform better in Test 3 than the other two groups. However, as previously discussed, no statistically detectable differences were found in Test 2 and Test 3 across three groups, it seems

unnecessary to compare participants' performances in Test 3 to detect if the transfer of effects on metacomprehension accuracy actually happens on the 4th reading. In this case, it is concluded that no transfer of effects on metacomprehension accuracy is observed in Test 3 across three groups.

4.2.2 The Analyses of the Time for Activities recorded in gStudy log files and Results

The log files recorded participants' activities in gStudy. Logged data were generated automatically in gStudy during the experiment sessions. Example activities include viewing the browser window, labelling texts and updating notes when participants interacted with learning materials in gStudy.

In this study, I was interested to examine possible differences in the time participants spent on various tasks including reading texts, tests, restudying and writing summaries among three groups. More specifically, the following time variables were defined: overall studying time, time spent on each of the four reading texts, time spent responding to items on Test 2 and Test 3, time spent on restudying the three reading texts, and for groups in which participants wrote summaries, time spent on writing summaries.

Skewness statistics for those variables were acceptable except for the variable Restudy, varying between .05 and 2.66. The variable Restudy has high positive skew 4.08 in the delayed-summary group, so a logarithmic (base 10) transformation was applied to reduce skewness (Tabachnick & Fidell, 2007). After the log₁₀ transformation, skewness was reduced from 4.08 to 1.38, and kurtosis was reduced from 16.36 to 2.18. The

Restudy variable is then renamed as “RestudyLog10” to be used in the latter data analysis.

Table 4.8 shows descriptive statistics for time participants spent on activities.

Table 4.8 Descriptive Statistics of Time of Activities in Three Groups

Group		N	Minimum	Maximum	Mean	Std. Deviation
No summary	All(sec)	24	1607.00	5290.00	2968.79	917.71
	Reading1(sec)	24	162.00	1072.00	366.67	200.01
	Reading2(sec)	24	129.00	640.00	279.42	118.89
	Reading3 (sec)	24	151.00	557.00	309.33	113.43
	Reading4 (sec)	24	99.00	995.00	389.92	167.42
	Test2(sec)	24	322.00	1151.00	678.00	204.98
	Test3(sec)	24	126.00	910.00	272.08	166.56
	Restudy(sec)	24	31.00	779.00	243.71	221.40
	ReStudyLog10	24	1.49E0	2.89E0	2.21E0	.42
	Valid N (listwise)		0			
Immediate Summary	All(sec)	21	2656.00	7318.00	4117.86	1160.44
	Reading1(sec)	21	187.00	768.00	448.67	190.43
	Reading2(sec)	21	208.00	666.00	372.29	134.48
	Reading3 (sec)	21	175.00	757.00	343.14	150.18
	Reading4 (sec)	21	116.00	692.00	355.24	152.28
	Test2(sec)	21	335.00	1046.00	586.14	197.46
	Test3(sec)	21	132.00	343.00	206.14	62.68
	Restudy(sec)	21	39.00	582.00	222.67	169.17
	Summary1 (sec)	21	217.00	912.00	412.95	173.23
	Summary2 (sec)	21	122.00	717.00	304.52	166.16
	Summary3 (sec)	21	110.00	1245.00	391.24	287.00
	ReStudyLog10	21	1.59E0	2.76E0	2.21E0	.37
	Valid N (listwise)		21			

Table 4.8 Descriptive Statistics of Time of Activities in Three Groups

Group		N	Minimum	Maximum	Mean	Std. Deviation
Delayed	All(sec)	24	1905.00	7503.00	4195.63	1482.84
Summary	Reading1(sec)	24	71.00	813.00	473.00	215.36
	Reading2(sec)	24	128.00	936.00	319.71	168.78
	Reading3 (sec)	24	135.00	562.00	324.29	125.63
	Reading4 (sec)	24	112.00	742.00	350.88	182.94
	Test1(sec)	24	160.00	719.00	401.96	164.81
	Test3(sec)	24	83.00	937.00	259.46	178.63
	Restudy(sec)	24	16.00	1920.00	268.29	382.16
	Summary1 (sec)	24	149.00	928.00	424.13	231.03
	Summary2 (sec)	24	105.00	773.00	286.63	169.87
	Summary3 (sec)	24	99.00	842.00	348.17	175.73
	ReStudyLog10	24	1.20E0	3.28E0	2.19E0	.46
	Valid N (listwise)	24				

Levene statistics testing homogeneity of variance in time variables across groups are presented in Table 4.9. Only the data distribution of All (overall study time) probably are not homogeneous ($p < .05$) across groups.

Table 4.10 shows a set of one-way ANOVAs comparing the mean time participants spent on various activities in the three groups. Given multiple tests and the failure to satisfy the assumption of homogeneity of variance for the overall time, it appears there are no statistically detectable differences across groups on any of these time variables.

Table 4.9 Test of Homogeneity of Variances on Activity Variables

Activities	Levene Statistic	df1	df2	Sig.
All(sec)	3.22	2	66	.05
Reading1(sec)	.95	2	66	.39
Reading2(sec)	.79	2	66	.46
Reading3 (sec)	.45	2	66	.64
Reading4 (sec)	.65	2	66	.53
Test2(sec)	.02	2	66	.98
Test3(sec)	2.67	2	66	.08
ReStudyLog10	.21	2	66	.81

Table 4.10 One-way ANOVA test for the Mean Time of Activities Spent in Three Groups

Activities		Sum of Squares	df	Mean Square	F	Sig.
All(sec)	Between Groups	2.23E7	2	1.11E7	7.58	.001
	Within Groups	9.69E7	66	1467811.70		
Reading1(sec)	Between Groups	147826.44	2	73913.22	1.80	.17
	Within Groups	2712092.00	66	41092.30		
Reading2(sec)	Between Groups	96741.53	2	48370.77	2.38	.10
	Within Groups	1341967.08	66	20332.84		
Reading3 (sec)	Between Groups	12813.05	2	6406.53	.38	.69
	Within Groups	1110030.86	66	16818.65		
Reading4 (sec)	Between Groups	21647.47	2	10823.74	.38	.69
	Within Groups	1878226.27	66	28457.97		
Test2(sec)	Between Groups	97684.72	2	48842.36	1.12	.33
	Within Groups	2875155.20	66	43562.96		
Test3(sec)	Between Groups	53853.84	2	26926.92	1.23	.30
	Within Groups	1450532.36	66	21977.76		
ReStudyLog10	Between Groups	.01	2	.003	.02	.99
	Within Groups	11.62	66	.18		

A post-hoc multiple comparison test investigating time differences for the entire study among three groups shown in Table 4.11 reports the Bonferroni multiple comparisons test result for the time spent in the study to find out how they differ from each other among three groups. It shows there are large and statistically detectable differences between Group 1 and Group 2 (1149 seconds) and between Group 1 and Group 3 (1226 seconds). There is no real difference for the time spent on the study between Group 2 and Group 3 (77 seconds).

Table 4.11 Multiple Comparisons Test for Time Spent on Activities in Three Groups
Bonferroni

Dependent Variable	(I) Groups	(J) Groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
All(sec)	No summary	Immediate summary	-1149.07*	362.02	.01	-2038.36	-259.77
		Delayed summary	-1226.83*	349.74	.002	-2085.98	-367.69
	No summary	Immediate summary	1149.07*	362.02	.01	259.77	2038.36
		Delayed summary	-77.77	362.02	1.00	-967.07	811.53

*. The mean difference is significant at the 0.05 level.

Table 4.12 shows descriptive statistics for time spent on writing summaries in the immediate and delayed groups.

Table 4.13 reports the results of the t-test comparing the mean values of the time spent on writing summaries in the immediate and delayed groups. The Levene's test for equality of variance indicates the assumption of homogeneity was met.

Table 4.12 Descriptive Statistics of the Mean of Time (Summaries) in Two Groups

Groups		N	Mean	Std. Deviation	Std. Error Mean
Summary1 (sec)	Immediate summary	21	412.95	173.23	37.80
	Delayed summary	24	424.13	231.03	47.16
Summary2 (sec)	Immediate summary	21	304.52	166.16	36.26
	Delayed summary	24	286.63	169.87	34.66
Summary3 (sec)	Immediate summary	21	391.24	287.00	62.63
	Delayed summary	24	348.17	175.73	35.87

Table 4.13 The t-test Result of the Mean of the Time (Summaries) in Two Groups

Summary		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Summary1 (sec)	Equal variances assumed	2.65	.11	-.18	43	.86	-11.17	61.61	-135.41	113.07
Summary2 (sec)	Equal variances assumed	.02	.88	.36	43	.72	17.90	50.25	-83.43	119.23
Summary3 (sec)	Equal variances assumed	1.38	.25	.62	43	.54	43.071	69.97	-98.03	184.18

There were no statistically detectable differences for the time spent writing summaries in the immediate and delayed groups.

4.2.3 The Analyses of Summaries for the Immediate and Delayed Groups and Results

In this study, participants who were assigned randomly to the immediate and delayed-summary groups were not instructed about how to write summaries. Some participants who inquired about the length of summaries were told that they could write to a length that they judged to sufficiently cover the content. Given the flexibility of writing the summaries, it is expected that summary data can greatly vary among participants. Meanwhile, time may indicate how much participants monitored and evaluated their summaries.

The summaries collected in this study were analyzed using a simplified rubric adapted from a scoring model developed by Schommer and Surber (1986). Summaries were decomposed into separate concepts and relations following a model originally introduced by Brown and Day (1983) and Kintsch (1979) that generates a “macrostructure” representation of a text. Each sentence in a summary was defined as consisting of concept units and relations. For example, in the sentence “customers care about product prices”, the concepts units are “customers” and “product prices,” and the verb “care” is their relation. Each summary generated in the experiment was scored against the model summary. Four types of summary errors were identified: concept units missed, concepts units distorted, relations missed and relations distorted. Concepts units or relations missed means that participants didn’t include major concepts and relations in

the reading texts in writing their summaries. Concepts units or relations distorted was defined as misunderstanding or misinterpreting the information presented in the reading texts or adding irrelevant information to the texts. All 4 types of errors were treated equally in a simplified rubric. That is, either for missing or distorting a component in the summaries, 1 score was subtracted from the full mark of the summaries. In Table 4.14, for Reading 1 and 2, all the components add up to 24 marks. For Reading 3, the total mark scores 18.

Table 4.14 Model Summaries in Reading 1, 2 and 3 in Two Groups

Model Summary	Reading1	Reading 2	Reading 3
Concept units	16	16	12
Relations	8	8	6
Total marks	24	24	18

Table 4.15 The Descriptive statistics of Summary Scores in Immediate and Delayed Groups

Groups		N	Minimum	Maximum	Mean	Std. Deviation
		Statistic	Statistic	Statistic	Statistic	Statistic
Immediate group	Summary1 scores	21	.00	21.00	6.8571	6.71033
	Summary 2 scores	21	.00	18.00	10.0952	4.07314
	Summary 3 scores	21	2.00	17.00	10.8095	4.10632
	Valid N (listwise)	21				
Delayed summary	Summary 1 scores	24	2.00	15.00	7.7083	3.88396
	Summary 2 scores	24	.00	18.00	9.1250	3.90443
	Summary 3 scores	24	6.00	16.00	11.0833	2.71736
	Valid N (listwise)	24				

Table 4.15 shows descriptive statistics for the summary scores that participants have obtained in the immediate and delayed groups. The standard deviations in two groups for summaries in each reading show there are great differences in the summary scores participants have got for the three reading texts.

Table 4.16 reports the results of the t-tests comparing the mean values of the summary scores between the immediate and delayed groups. The Levene's test for equality of variance showed that equal variance could be assumed for the t-tests ($p < 0.05$) except for data describing Summary 2.

As can be seen in Table 4.16, the mean summary scores across the two groups were not statistically detectably different ($p > .20$).

Table 4.16 The t-test Result of the Mean for the Immediate and Delayed Groups

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Summary1 scores	Equal variances assumed	4.00	.05	-.53	43	.60	-.85	1.61	-4.10	2.39
Summary2 scores	Equal variances not assumed	.12	.73	.81	41.67	.42	.97	1.19	-1.44	3.38
Summary3 scores	Equal variances assumed	4.9	.03	-.27	43	.79	-.27	1.03	-2.34	1.80

4.2.4 The Analyses of Questionnaires and Results

The questionnaire used in this study collected demographic information, such as age, major and the native language they speak, and participants' reflections on features of the study: which texts they reported restudying, the number of texts they reported restudying, their attitude toward using summaries to judge their understanding and using restudy as a learning strategy. The data from the same 69 participants as have been considered so far were analyzed.

In the questionnaire, participants reported the types of information they restudied in this experiment as well as in general. Specifically, in question 5, participants described kinds of information restudied in the three texts. Question 6 inquired about kinds of information participants usually restudy in their courses. Table 4.17 and Table 4.18 show the frequency of participants' responses for questions. For question 5, the item participants in three groups chose most was "the part(s) about which I was confused" as what they restudied in the experiment. There were two participants in the three groups who indicated they did not restudy any part of the three reading texts. For question 6, the item that participants chose most was "the part(s) about which I am confused" to describe their general restudy targets in courses. There were three participants who indicated that they never reread a text unless asked to do so.

Table 4.17 Participants' opinions on Restudy Areas in Q5

Group	Q5-a The part(s) that in which I am most interested	Q5-b The part(s) that I think are difficult	Q5-c The part(s) about which I was confused	Q5-d The part(s) I thought I might not recall for the tests	Q5-e I just reread the whole article again	Q5-f None of the above describes what I restudied.	Q5-g I did not restudy any part(s) of the 4 readings
No summary	42.9%	36%	62.5%	44%	12.5%	4.2%	0%
Immediate summary	23.8%	42.9%	71.4%	57.1%	9.5%	0%	0%
Delayed summary	12.5%	42.9%	54.2%	62.5%	12%	0%	8.3%

Table 4.18 Participants' opinions on Restudy Areas in Q6

Group	Q6-a The part(s) that I am the most interested in	Q6-b The part(s) that I think very difficult	Q6-c The part(s) that I feel very confused	Q6-d The part(s) that I feel I cannot recall in the future test	Q6-e I only do reread occasionally while studying	Q6-f I never reread a reading unless I am asked to do so
No summary	21.8%	62.5%	79.1%	45.8%	12.5%	12.5%
Immediate summary	19%	76.2%	85.7%	52.4%	9.5%	0%
Delayed summary	21.8%	67.8%	79.1%	62.5%	12.5%	12.5%

Table 4.19 shows participants' opinions on the usefulness of writing summaries and restudying. Questions Q7 and Q10 focus on participants' opinions about the usefulness of restudy and the likelihood of using it in the future. The 5 responses (strongly agree, agree, neutral, disagree, strongly disagree) were re-indexed into numbers 5, 4, 3, 2, and 1. The scores were then averaged by group.

Table 4.19 Restudy and Summary Attitude in the Three Groups

Group		N	Mean	Std. Deviation
No summary	Q7 scores	24	4.21	.72
	Q10 scores	24	3.88	.80
	Valid N (listwise)	0		
Immediate summary	Q7 scores	21	4.33	.66
	Q10 scores	21	4.29	.78
	Q8 scores	21	4.05	.92
	Q9 scores	21	3.67	.86
	Q11 scores	21	3.57	1.08
	Valid N (listwise)	21		
Delayed summary	Q7 scores	24	4.13	.68
	Q10 scores	24	3.62	1.25
	Q8 scores	24	3.92	1.02
	Q9 scores	24	3.21	1.35
	Q11 scores	24	3.46	1.14
	Valid N (listwise)	24		

Table 4.20 reports the homogeneity test for Q7 and Q10. It shows that homogeneity of variances is satisfied. In Table 4.21, a one-way ANOVA compares group differences on their opinions towards restudy on these 2 questions. There were no

statistically detectable differences for their opinions on the usefulness of restudy in the current study. As it is indicated in Table 4.19, the mean answers to those 2 questions in all three groups are all approximately 4 (agree). This means all participants across groups agreed that restudy was helpful for them in the study and they will be likely to use it in their future studying.

Table 4.20 Test of Homogeneity of Variances on Q7 and Q9

	Levene Statistic	df1	df2	Sig.
Q7 scores	1.297	2	66	.280
Q10 scores	2.301	2	66	.108

Table 4.21 One-way ANOVA Test for Opinions of Q7 and Q10 in Groups

		Sum of Squares	df	Mean Square	F	Sig.
Q7 scores	Between Groups	.591	2	.295	.626	.538
	Within Groups	31.148	66	.472		
	Total	31.739	68			
Q10 scores	Between Groups	5.465	2	2.732	2.871	.064
	Within Groups	62.825	66	.952		
	Total	68.290	68			

In the questionnaire, participants in the immediate and delayed summary groups were asked their opinions about using summaries to judge their understanding and to direct them to restudy in three questions. Questions Q8, Q9, and Q11 indicate

participants' opinions on the usefulness of writing summaries and the likeness of using them in the future study.

Table 4.22 T-test for Comparing Opinions of Q8, Q9 and Q11 Across Groups

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Q8 Equal scores variances not assumed	.063	.803	.631	42.174	.531	.180	.285	-.395	.755
Q9 Equal scores variances not assumed	1.548	.220	1.241	41.594	.222	.410	.330	-.257	1.077
Q11 Equal scores variances not assumed	.826	.368	.237	41.757	.814	.080	.337	-.601	.761

Table 22 reports the results of t-tests comparing opinions about using summaries to judge understanding of texts in between the immediate and delayed summary groups. The Levene's test for equality of variance showed that equal variance could not be assumed for all three questions among the two groups. There were no statistically detectable differences in participants' opinions on the usefulness of using summaries to judge their understanding in the current context or in future studying.

More specifically, in both immediate-summary ($M=4.05$) and delayed-summary groups ($M=3.92$), participants agreed that summarizing helps them evaluate how well

they understand the texts. Participants in both the immediate-summary group (M=3.67) and in the delayed-summary group (M=3.21) agreed that summarizing helps them to plan what to restudy. Participants in both the immediate-summary (M=3.57) and delayed-summary groups (M=3.46) agreed it is likely they will use summarization to how well they understand what they study.

CHAPTER 5: DISCUSSIONS

5.1 Results Discussions

5.1.1 Metacomprehension Accuracy

In the current study, metacomprehension accuracy was defined as the correlation between comprehension judgments and test performances. Because of the unacceptable low reliability of Test 1, relations between judgments of learning and test performance could not be investigated in that context. However, ratings of understanding across the three groups were examined. As participants were randomly assigned to one of the three groups, they were assumed to be equivalent in their comprehension abilities prior to participating in the study. The one-way ANOVA reported in chapter 4 indicated that no statistically detectable differences were found among the three groups on comprehension judgments. This means that participants were similar in judging their understanding – their metacomprehension accuracy did not vary across the three conditions of no summaries, immediate summaries and delayed summaries.

5.1.2 The Test Performances among the Three Groups

As previously discussed, metacomprehension accuracy concerns how well a text is understood. It is hypothesized that the higher the metacomprehension accuracy, the more effective a learner can regulate his/her study and thus the higher the test performance. As a consequence of no differences observed in metacomprehension accuracy among the three groups, according to the hypothesis, it is not surprising to find

there were no statistically detectable differences in test performance in Test 2 among the three groups. Likewise, there also were no statistically detectable differences among the three groups for Test 3 covering the 4th reading, which was intended to test the transfer of the effect of repeated opportunity to observe relations between judgments of learning and restudying. Writing either immediate or delayed summaries had no impact on participants' reading comprehension.

These results contradict the findings of studies that Anderson and Thiede (2003, 2008) conducted to examine the effects of writing delayed summaries on improving learners' metacomprehension accuracy. In their studies, metacomprehension accuracy was operationalized by the correlation between the judgment made after writing delayed summaries and test performance. They found differences in test performance favoring the delayed-summary group compared to the immediate-summary and no-summary group. As the experiment design in the current study is similar to Anderson and Thiede's (2003, 2008), plausible reasons that could account for the contradictory results are discussed in the next section.

5.1.3 The Quality of Summaries Written in the Immediate and Delayed Groups

The analysis of the quality of summaries in the study showed how well participants composed summaries in the immediate and delayed summary groups. In this study, participants were not instructed how to write summaries and nor told there was an expected length for the summaries. They were only told to write summaries that they judged to be complete and as accurate as possible within a reasonable length. Thus, the quality of the summaries not only provides a valid clue of how well participants have studied the reading texts, but also indicates how well participants metacognitively judged

the quality of their summaries. Presumably, summaries may have different validity in directing restudy.

In both immediate-summary and delayed-summary groups, participants wrote the summaries before they rated their understanding of each reading. When writing a summary, a learner retrieves available information from memory and connects propositions to form the summary. By this means, participants have access to an impression of how well they understood the reading texts based on how easy it was to compose the summaries and how good they were. Thus, it is inferred that participants used the quality of their summaries, including the quantity of information retrieved and included in the summaries, as a cue to estimate their understanding of the reading texts. However, no statistically detectable differences between the two groups were found for the quality of summaries. This corresponds to the finding of no differences in ratings of understandings. Because ratings of comprehension did not differ among the groups and summaries did not differ among groups either, it would not be surprising that there would be no difference in how these participants restudied.

5.1.4 The Effects of Restudy on Achievement

In the current study, participants were provided opportunities to restudy the three texts before they proceeded to the 2nd test. As it was originally predicted that learners would have different metacomprehension accuracy by writing no, immediate or delayed summaries, they were expected to perform differently during the restudy session, and on Test 2 and Test 3.

The effect of restudying can be observed from group differences in participants' test performances on Test 2. As previously discussed, there were no statistically detectable differences found among groups' performance on Test 2. This is consistent with the no observed differences on participants' ratings on the reading texts. Thus, it is concluded that restudying did not differ and consequently test scores did not differ among the three groups as predicted.

Although restudy time cannot be used as a direct indicator of how participants studied in the restudy period, it provides a general index of judged need to restudy. There were no statistically detectable differences in restudy time across groups. This may reflect similarities among the participants in the three groups about how they regulated restudying even after receiving different treatments in the study.

5.1.5 Attitudes toward Using Summaries and Restudy as Strategy in the Current Study and in the Future

In the current study, participants completed a self-report questionnaire about their experiences in the study and their use of the strategy of writing summaries and restudying. All participants in the three groups agreed that restudy helps them learn better in the current experiment. They also suggested they are likely to use it to judge how well they have studied on other occasions.

Participants in both the immediate-summary and delayed-summary groups shared similar opinions about the effectiveness of writing summaries as cues to direct their learning activities. For writing summaries, participants in both groups agreed that they found it helps them to evaluate how well they comprehended texts in the current study. They reported they will use summarizing when they study in the future. They also

expressed the view that summarizing texts helps them to plan what to restudy in the current study and is likely to guide their future study.

5.2 Conjectures for Results and Possible Explanations

The results reported in this study failed to ratify the delayed-summary effects found in previous research. In this section, I consider five conjectures that offer potential explanations for these results and their validity.

5.2.1 Did Participants Spend Sufficient Time Studying and Taking Tests?

One conjecture to explain the observation of no differences among the three groups may be associated with insufficient time participants spent on reading the texts and/or taking the tests. If a participant spends insufficient time studying, he or she may not be able to perform well. Likewise, if a participant takes less than a reasonable amount of time to answer test items, they likely cannot do well in the test.

However, several reasons can be provided to justify that this conjecture is invalid. First, participants were notified they could take as long as they needed to study the texts. Also, they had the opportunity to restudy each text with no time constraint and allowing an unlimited number of visits. There was also no time limit for taking the tests. Therefore, participants would not experience time pressure to finish these activities. Were data about durations valid and did they reflect participants' real levels of performances in the study? Yes. The durations for all the activities in the study were recorded for each participant in the log file generated in the gStudy program. By setting a minimum reading and answering time as discussed in chapter 4, based on estimates of

reasonable speeds, some participants were removed from the data pool because it was judged they had not dedicated reasonable effort to tasks.

5.2.2 Were the Reading Materials Too Difficult for Participants to Understand?

The second conjecture concerns the difficulty level of the reading texts used in the study. It is sensible that if the reading materials were too unfamiliar or obscure for participants, metacomprehension accuracy would not vary across groups because it would be uniformly low. In the current study, the difficulty level of the reading texts was estimated using the Flesch-Kincaid index (Flesch, 1948). Similar to the reading texts in Anderson and Thiede's studies (2003, 2008), the readability of the texts ranges from 9.3 to 12.2, with an average Flesch-Kincaid Grade Level of 10.6. These levels seem appropriate for an undergraduate population.

However, as Anderson and Thiede acknowledged, readability scores may not really reflect text difficulty. Unfortunately, as they did not provide further difficulty measurements for the reading texts in their study, it is not feasible to otherwise compare the difficulty levels of the reading texts in their studies and mine, thus preventing me from identifying whether differences in the difficulty of texts used the two studies could account for the variant findings.

How did participants rate their understanding of the texts? These scores could range from 0 to 10. The average rating aggregated over all participants was 7.20 (Reading 1), 7.57 (Reading 2), and 7.96 (Reading 3), respectively. This indicates that participants in general thought that they understood the articles well. However, as ratings are reported by participants themselves, they are not an objective evaluation of the actual difficulty of

the reading texts. Nonetheless, the readability scores for texts and participants' self-generated ratings together imply the texts were not particularly difficult.

5.2.3 Were the Tests Too Difficult?

Similarly, the difficulty level of tests in the study was also examined to explore if this factor might lead to the unexpected results of this study. If the tests were too difficult, any effect of metacomprehension accuracy across groups would also be expected to go undetected. Consider the general performance level aggregated over all participants.

As reported in Chapter 4, the average scores of Test 2 and Test 3 for aggregated across the three groups were 13.1/27 or 48.5%, and 5.62/9 or 62.4%, respectively. While these values are less than would be expected in many classrooms, they are ideal for revealing variances due to individual differences and potential treatment effects.

Such an analysis may lead us to conclude that it is plausible that metacomprehension accuracy only matters when the difficulty of reading texts or tests is at a different level for learners. If a text is more challenging than is typical, even with a high level of metacomprehension accuracy, learners still cannot achieve high test performance because it is beyond their capabilities to cope with it.

5.2.4 Metacomprehension Accuracy Needs Practice

Although the failure to find effects for writing delayed summaries may be due to any of the foregoing factors, it also may be that, like mastering other kinds of skills, learners need to (a) learn metacomprehension skills and (b) practice their

metacomprehension monitoring skills before they can freely deploy them effectively. In the current experiment, participants were assumed to have well developed metacomprehension monitoring skills. So, when they wrote immediate or delayed summaries, they would be able to use these skills effectively to evaluate accurately whether they understood an text well. They were also assumed to be able to regulate restudy accordingly to achieve high performance. However, it is possible that participants may not be able to extract appropriate metacomprehension information nor properly use it to regulate studying.

In fact, other research has confirmed that metacomprehension skills are not well developed among undergraduate and other readers who monitor their understanding (e.g. Wiley, Griffin, & Thiede, 2005; Glenberg, Wilkinson & Epstein, 1982; Maki & Berry, 1984; Weaver, 1990). This suggests learners may not be aware that how well they have written summaries is an indicator of the completeness and accuracy of their comprehension of what they have read.

Moreover, even if learners have basic metacomprehension skills, they still may need to practice these skills before they can accurately employ them in learning. Findings in Vesonder and Voss's study (1985) on learners' prediction accuracy demonstrated that prediction accuracy improves after learners have practiced their monitoring skills. In the current study, learners may not have been able to polish predictions of their performance due to low metacomprehension skills. Because metacomprehension accuracy influences learning outcome through its affects on regulation of study (Anderson & Thiede, 2008), participants may not have acquired higher metacomprehension accuracy and thus failed

to better regulate their studies during the restudy sessions, resulting in no performance differences among the other groups.

In this sense, learners may need more sessions than were provided before they can benefit from the effects of writing delayed summaries and strategically apply such knowledge to regulate their study and restudy.

CHAPTER 6: CONCLUSION AND FUTURE WORK

6.1 Conclusion

A review of research on improving learners' metacomprehension accuracy showed that writing delayed summaries after reading helps improve learners' metacomprehension accuracy and may influence their regulation of future studying which may lead to differentiated learning outcomes. To explore the validity of this claim, I compared the effects of writing delayed summaries to writing no summaries or immediate summaries. Results did not align with theory or previous research – no statistically detectable differences were found across groups. Conjectures about why such results were observed were suggested in terms of learners' text difficulties, test difficulties and inadequate practice on metacomprehension monitoring.

One of the implications of the current study is that the difficulty of information to be learned influences learners' performance in relation to the availability and skillfulness of tactics for learning. If texts are too difficult for learners, metacomprehension accuracy can have little effect on achievement if the learner lacks study tactics that can repair gaps in knowledge or otherwise compensate. Thus, providing learning materials at an appropriate level of difficulty, in the sense of matched to opportunities for learners to productively use study tactics they have, may be a key to promoting achievement.

Another consideration regarding improving metacomprehension accuracy is that learners may need practice with metacomprehension skills before they can flexibly apply them in monitoring their progress. In previous studies, learners were assumed capable of

monitoring their progress productively as long as they have provided the correct cues prior to their judgment. However, due to the inferential nature of judgments, learners may need still need practice in judging metacomprehension to reach a sensible conclusion about their study progress.

It is concluded that the process of improving metacomprehension accuracy is a more complex procedure than has been acknowledged. The manipulation of metacomprehension requires taking a consideration of several new potential factors involved in each phase of learners' regulation of study.

6.2 Limitation and Future Research Work

Like other types of skills, metacomprehension skills take time to develop. One limitation of the current study is that participants joined in only one experiment session and may not have had sufficient time to practice those skills to achieve a high level of metacomprehension accuracy. Ideally, participants would participate in a series of training sessions, until they understand how to monitor their understanding of texts and have skill to then apply metacomprehension tactics to regulate studying effectively. Consequently, lack of adequate practice may have contributed to not observing the effect of writing delayed summaries found in previous research studies by Anderson and Thiede (2003, 2008). Thus, in the future study, it is recommended that sufficient practice be provided to build learners' metacomprehension skills.

Restudy is an indispensable component in learners' regulation of study. Without proper enactment of learning tactics during restudy, learners can't perform well even with

a high level of metacomprehension accuracy on texts. In the current research, some but perhaps not sufficient data about how participants restudied are recorded in the gStudy-generated logging files. Given the crucial role of restudy plays as an expression of how learners regulate learning, future research should “look inside” the black box of what learners do when they restudy to identify proximal variables that affect learning.

Another limitation of the current study is that the first test and its subscales defined in terms of cognitive complexity could not be used due to its low reliability. Analyses planned that used those data could have provided straightforward evidence to differentiate levels of metacomprehension accuracy across treatment conditions and types of tasks on the achievement test. Future studies should avoid this problem.

Although there appear to be some successes in improving metacomprehension accuracy in prior experiments, this experiment demonstrates that techniques such as writing delayed summaries or judging one’s study in a delayed fashion do not guarantee high levels of metacomprehension accuracy or reading comprehension (Dunlosky & Lipko, 2005). Therefore, discovering methods that can produce a *consistently* high level of metacomprehension accuracy remains central goal for future research on metacomprehension accuracy.

REFERENCE LIST

- Azevedo, R., & Hadwin, A. F. (2005). Scaffolding Self-regulated Learning and Metacognition - Implications for the Design of Computer-based Scaffolds. *Instructional Science*, 33, 367-379.
- Armbruster, B. B., Anderson, T. H., & Ostertag, J. (1987). Does text structure/summarization instruction facilitate learning from expository text? *Reading Research Quarterly*, 22, 331–346.
- Anderson, L.W., & Krathwohl, D. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Anderson, M. C. M., Thiede, K. W. (2008). Why do delayed summaries improve metacomprehension accuracy? *Acta Psychologica*, 128, 110-118
- Azevedo, R. (2005). Using Hypermedia as a Metacognitive Tool for Enhancing Student Learning? The Role of Self-Regulated Learning. *Educational Psychologist*, 40(4), 199-209.
- Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, 127, 55 - 68.
- Brown, A. L., & Day, J. D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior*, 22, 1-14.
- Britton, B. K., & Gülgöz, S. (1991). Using Kintsch's model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*, 83, 329-345.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65, 245–281.
- Caroll, M., & Nelson, T.O. (1993). Effect of overlearning on the feeling of knowing is more detectable in within-subject than in between-subject designs. *American Journal of Psychology*, 106, 227-235
- Corno, L., & Mandinich, E. B. (1983). The role of cognitive engagement in classroom learning and motivation. *Educational Psychologist*, 18, 88– 108.
- Dabbagh, N. (2003). Scaffolding: an important teacher competency in online learning. *TechTrends for Leaders in Education and Training*, 47(2), 39-44

- Dabbagh, N., & Kitsantas, A. (2005). Using Web-based Pedagogical Tools as Scaffolds for Self-regulated Learning. *Instructional Science*, 33, 513-540.
- Dabbagh, N. & Bannan-Ritland, B. (2005). Online learning: Concepts, strategies, and application. Upper Saddle River, NJ: Prentice Hall, Inc.
- Doctorow, M. J., Wittrock, M. C., & Marks, C. B. (1978). Generative processes in reading comprehension. *Journal of Educational Psychology*, 70, 109-118.
- Dunlosky, J., & Lipko, A. R. (2005). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science*, 16, 228-232.
- Dunlosky, J. , & Nelson, T.O. (1992). Importance of the kind of cue for judgment for learning (JOL) and the delayed OL effect. *Memory & Cognition*, 20, 374-380
- Dunlosky, J., Rawson, K.A., & Middleton, E.L.(2005). What Constrains the Accuracy of Metacomprehension Judgments? Testing the Transfer-Appropriate-Monitoring and Accessibility Hypotheses. *Journal of Memory and Language*, 52(4), 551-565
- Fletcher, C. R., van den Broek, P., & Arthur, E. J. (1996). A model of narrative comprehension and recall. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 141–164). Mahwah, NJ: Erlbaum.
- Friend, R. (2001). Effects of strategy instruction on summary writing of college students. *Contemporary Educational Psychology*, 26(1), 3–24
- Griffin, T. D., Wiley, J., & Thiede, K.W. (2007). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition*, 36, 93-103.
- Gunawardena, C. N., & Zittle, F. J (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *The American Journal of Distance Education*, 11(3), 8– 26.
- Hartley, K. & Bendixen, L. D. (2003). The use of comprehension aids in a hypermedia environment: Investigating the impact of metacognitive awareness and epistemological beliefs. *Journal of Educational Multimedia and Hypermedia*, 12(3), 275-289
- Hiltz, S. R. (1997). Impacts of college-level courses via asynchronous learning networks: Some preliminary results. *Journal of Asynchronous Learning Networks*, 2. Retrieved November 2, 2008 at http://www.aln.org/publications/jaln/v1n2/pdf/v1n2_hiltz.pdf.
- Kauffman, D. F. (2004). Self-regulated learning in Web-based environments: Instructional tools designed to facilitate cognitive strategy use, metacognitive processing, and motivational beliefs. *Journal of Educational Computing Research*, 30, 139–161.
- Kimball, D. R., & Metcalfe, J. (2003). Delaying judgments of learning affects memory, not metamemory. *Memory & Cognition*, 31, 918-929.

- Kintsch, W. & Van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85 (5), 363-394.
- Kintsch, W. (1979). On modeling comprehension, *Educational Psychologist*, 14, 3-14.
- Kintsch, W. (1998) Comprehension: A paradigm for cognition. New York: Cambridge University Press.
- Leonesio R.J., & Nelson T.O. (1990) Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning Memory and Cognition*, 16, 464-470.
- Ley, K., & Young, D. (2001). Instructional principles for self-regulated learning. *Educational Technology Research and Development*, 49(2), 93-102.
- Lovelace, EA. (1984a). Metamemory Monitoring future recallability during study. *Journal of Experimental Psychology Learning Memory and Cognition*, 10, 756-766.
- Maki, R. H. (1998b). Test predictions over text material. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 117-144). Mahwah, NJ: Erlbaum.
- Maki, R. H., Jonas, D. & Kallod, M. (1994). The relationship between comprehension and metacomprehension ability. *Psychonomic Bulletin & Review*, 1, 126-129.
- Maki, R. H., Shields, M., Wheeler, A. E., & Zacchilli, T. L. (2005). Individual differences in absolute and relative metacomprehension accuracy. *Journal of Educational Psychology*, 97, 723-731.
- Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs? *Memory & Cognition*, 29, 222-233
- McLoughlin, C., Winnips, K. & Oliver, R. (2000). Supporting Constructivist Learning through Learner Support On-line. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2000* (pp. 674-680). Chesapeake, VA: AACE.
- Metcalf, J. (2002). Is study time allocated selectively to a region of proximal learning? *Journal of Experimental Psychology: General*, 131, 349-363.
- Morris, C. C. (1990). Retrieval processes underlying confidence in comprehension judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 223-232.
- Narciss, S., Proske, A., & Koerndle, H. (2007/5). Promoting self-regulated learning in web-based learning environments. *Computers in Human Behavior*, 23(3), 1126-1144.

- Nelson, T.O., & Dunlosky J (1991). When people's judgments of learning (JOL) are extremely accurate at predict subsequent recall: The delayed-JOL effect. *Psychological Science*, 2, 267-270
- Nelson, T.O. and Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G.H. Bower (Ed.), *The psychology of learning and motivation* (pp. 125-141). New York: Academic Press.
- Niemi, H., Nevgi, A., & Virtanen, P. (2003). *Towards self-regulation in Web-based learning. Journal of Educational Media*, 28 (1), 49-71.
- Pearson, P. D., & Fielding, L. (1996). Comprehension instruction. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. II, pp. 815–860). Mahwah, NJ: Erlbaum.
- Pintrich, P. R., and de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33–40.
- Radmacher, S., & Latosi-Sawin, E. (1995). Summary writing: A tool to improve student comprehension and writing in psychology. *Teaching of Psychology*, 22, 113-15
- Rawson, K. A., & Dunlosky, J. (2002). Are performance predictions for text based on ease of processing? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 69-80.
- Rawson, K. A., Dunlosky, J., & Thiede, K. W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory & Cognition*, 28, 1004-1010.
- Richardson, J., & Swan, K. (2003). Examining social presence in online courses in relation to students' perceived learning and satisfaction, *Journal of Asynchronous Learning Networks*, 7(1), 68– 88.
- Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (2001). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education*, 14(2), 50– 71.
- Schommer, M & Surber, J. R. (1986). Comprehension-monitoring failure in skilled adult readers. *Journal of Educational Psychology*, 78(5), 353-357.
- Schunk, D.H. (1989) Social cognitive theory and self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed., pp. 153-189). New York: Longman.
- Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist*, 25, 71–86.

- Shaughnessy, J. J., & Zechmeister, E. B. (1992). Memory monitoring accuracy as influenced by the distribution of retrieval practice. *Bulletin of the Psychonomic Society*, 30, 125–128.
- Son, L. K., & Metcalfe, J. (2005). Judgments of Learning: Evidence for a Two-Stage Model. *Memory & Cognition*, 33, 1116-1129.
- Spellman, B. A., & Bjork, R. A. (1992). When predictions create reality: Judgments of learning may alter what they are intended to assess. *Psychological Science*, 3, 315-316.
- Stone, N. J. (2000). Exploring the Relationship between Calibration and Self-Regulated Learning. *Educational Psychology Review*, 12, 437-475.
- Swan, K. (2002). Immediacy, social presence, and asynchronous discussion. In J. Bourne & J. C. Moore (Eds.), *Elements of quality online education*, 3, Olin & Babson Colleges: Sloan Center for Online Education.
- Swan, K. (2003). Developing social presence in online course discussions. In S. Naidu (Ed.), *Learning and teaching with technology: Principles and practices*, 147–164. London: Kogan Page.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). New York: Pearson.
- Thiede, K. W., & Anderson, C. M. (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology*, 28, 129-160.
- Thiede, K. W., Dunlosky, J., Griffin, T. D., & Wiley, J. (2005). Understanding the Delayed-Keyword Effect on Metacomprehension Accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(6), 1267-1280.
- Uvelace EA (1984b) Metamemory Monitoring future recallability in free and cued recall *Bulletin of the Psychonomic Society* 22, 497-500
- van den Broek, P., Ridsen, K., Fletcher, C. R., & Thurlow, R. (1996). A “landscape” view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 165–187). Mahwah, NJ: Erlbaum.
- Van den Boom, G., Paas, F., van Merriënboer, J. J. G., & van Gog, T. (2004/7). Reflection prompts and tutor feedback in a web-based learning environment: effects on students' self-regulated learning competence. *Computers in Human Behavior*, 20(4), 551-567.
- Van Overschelde, J.P., & Nelson, T.O. (2006). Delayed judgments of learning cause both a decrease in absolute accuracy (calibration) and an increase in relative accuracy (resolution), *Memory and Cognition*, 34 (7), 1527-1538

- Weaver, C. A. III, & Kelemen, W.L. (1997). Judgments of learning at delays: shifts in response patterns or increased metamemory accuracy? *Psychological Science*, 8, 318-321
- Weaver, C. A., & Kelemen, W. L. (2003). Processing similarity does not improve metamemory: Evidence against transfer-appropriate-monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1058–1065.
- Whipp, J. & Chiarelli, S. (2004). Self-regulation in a web-based course: A case study. *Educational Technology Research and Development*, 52 (4), 5-22.
- Wiley, J., Griffin, T. D., & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *Journal of General Psychology*. 132(4), 408-428.
- Winne, P.H. (1995a). Inherent details in self-regulated learning. *Educational Psychologist*, 30 (4), 173-187
- Winne, P.H. (1995b). Self regulation is ubiquitous but its forms vary with knowledge. *Educational Psychologist*, 30 (4), 223–228
- Winne, P.H. (2005). A perspective on state-of-the-art research on self-regulated learning. *Instructional Science*, 33, 559–565.
- Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences*, 8, 327-353.
- Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed., pp. 153-189). New York: Longman.
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A.C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277– 304). Hillsdale, NJ: Erlbaum.
- Winne, P. H. & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds), *Handbook of Self-Regulation* (pp. 531-566). Orlando, Florida: Academic.
- Winograd, P.N. (1984). Strategic difficulties in summarizing text. *Reading Research Quarterly*, 19, 404-425.
- Wittrock, M. C., & Alesandrini, K. (1990). Generation of summaries and analogies and analytic and holistic abilities. *American Educational Research Journal*, 27, 489–502.
- Wood, D., Bruner, J., & Ross, G.(1976). The role of tutoring in problem solving. *Journal of Child Psychology and psychiatry*, 17, 89-100.
- Zimmerman, B. J. (1986). Development of self-regulated learning: Which are the key sub-processes? *Contemporary Educational Psychology*, 16, 307-313.

- Zimmerman, B. J. (1995). Self-regulation involves more than metacognition: A social cognitive perspective. *Educational Psychologist*, 30, 217-221.
- Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk & B.J. Zimmerman (Eds.), *Self-regulated learning: From teaching to self-reflective practice* (pp. 1–19). New York: Guilford.
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & Moshe Zeidner (Eds.), *Handbook of self-regulation* (pp. 13– 39). New York: Academic Press.
- Zimmerman, B.J. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed., pp. 153-189). New York: Longman.
- Zimmerman, B. J., & Tsikalas, K. E. (2005). Can computer-based learning environments (CBLEs) be used as self-regulatory tools to enhance learning? *Educational Psychologist*, 40, 267-271.

APPENDICES

Appendix 1

Four Texts Used in the Study

Text 1: Bearing Bad News

<http://www.nytimes.com/2008/06/15/books/review/Chotiner2-t.html?scp=24&sq=review&st=nyt>

FINAL SALUTE

A Story of Unfinished Lives.

In Steven Spielberg's "Saving Private Ryan," soon after the re-creation of D-Day, attention turns stateside to a middle-aged woman washing dishes as a car approaches her home. The vehicle comes to a stop, and the woman — knowing what awaits — steps outside and onto her porch before falling to the ground, almost as if she were a soldier wounded in battle. Spielberg chooses to place the camera behind his actress and at somewhat of a distance; we do not see her face when she hears the news that three of her sons have been killed in combat. There is no audible dialogue, either. The power of the scene lies in its universality: a parent collapsing from an emotional anguish projected over a barren landscape.

Jim Sheeler's "Final Salute" — inspired by a Pulitzer Prize-winning article he wrote for *The Rocky Mountain News* — tells the story of Major Steve Beck of the United States Marine Corps, whose job it is to deliver the news to the families of marines who have died. In this book, Sheeler reports on the wrenching and commonplace details of "casualty notification" — the moments that Spielberg's film leaves to the imagination. Here Beck and a Navy chaplain, on a typical day, inform a young woman that her husband has been killed in Iraq (two people are required for every such visit, Sheeler tells us, for physical protection as well as emotional support): "Finally she stood, but she still couldn't speak. As the major and the chaplain remained on their feet, she glared at them. It was a stare the major had seen before, the one that hurts the most."

Sheeler's book follows several families, from the time they receive notification and the burials of their loved ones to their attempts to find solace and move forward. In Major Beck, the author has found an ideal protagonist, someone who is simultaneously guarded and insightful. Like the families he must face, Beck finds comfort in the rituals and rules that guide the process. The corps has specific instructions for how to go about the job, but

the major has his own routines, too: “Beck made a point of learning each dead marine’s name and nickname. He touched the toys they grew up with and read the letters they wrote home.” Still, Beck soon understands why “even the ‘grunts’ on the front lines” say they prefer mortal danger to his job. “I can’t help but feel that I’m the person who’s bringing them all that pain,” he says to Sheeler. “That because I’m standing in front of them, they’re feeling as bad as they’re ever going to feel.”

War may indeed be an extension of politics, but Sheeler is intent on letting his narrative unfold in the most apolitical manner possible. He does note the heavy strain of casualties on an often underprepared military — one grieving father is shown his son’s coffin for the first time while it sits on a forklift — but Sheeler grasps that everyone reading his book will have preconceived, if nuanced, notions about America’s current conflicts, and therefore he can tell his story straight. The problem with “Final Salute” is that it reads like a series of newspaper articles strung together somewhat incoherently. And the episodic nature of the narrative may leave readers confused as to which story Sheeler is focusing on at any given time. (He should also get a prize for one of the most oddly arrogant insertions in modern memory. In an author’s note, he writes: “An editor at the newspaper pointed out a crucial similarity she said she saw in both Beck and myself: *You have caring in common.*”)

The book’s crucial flaw, however, is unintentionally revealed in an early remark Sheeler makes about Beck’s work: “While each door is different,” he writes, “the scenes inside are almost always the same.” Precisely because these scenes are so depressingly similar, the book feels padded as well as slight. Still, there is something oddly comforting about the nearly interchangeable reactions the families display when confronted with a crushing loss. Perhaps unwittingly, then, Sheeler has written something that captures the universal as much as it does the personal.

Text 2: Microsoft Should Know Money Can’t Buy Love

<http://bits.blogs.nytimes.com/2008/05/22/microsoft-should-know-money-cant-buy-love/>

The rap on Microsoft is that it copies everyone else’s good ideas. That may be too kind. I think it also copies a lot of proven bad ideas. To wit, Microsoft’s latest move to offer shopping rebates to people who use its Internet search service.

You can see why Microsoft would want to do this: It has too much money and not enough customers, so why not pay people to use its services? Sometimes putting money into creating value for your customers is a great strategy. For an example, see Amazon's decision to subsidize free shipping rather than buy television ads.

But Microsoft's approach is wrongheaded. First, it's not really about the main market for search. Price comparison shopping is a good service, but a relatively modest market, with none of the independent sites — like eBay's Shopping.com — or the services of search engines — like Google's Product Search — really taking off.

More importantly, history has proven that you can't build a mass Internet business by buying customers with cash or rewards. I've seen hundreds of companies that have tried versions of this. There was AllAdvantage, a company that wanted to pay people to surf the Web and see ads. There are all kinds of bonus and reward programs, like the now-defunct Netcentives, which offered airline miles for online shopping, and Upromise, which takes a cut of purchases and kicks it into a college savings account.

It appears that somewhat more people are motivated by social status and approval: look at the behavior on social networks and on some sites that create indicators of status for active members like Yelp. But far and away the biggest motivator of behavior online has been speed and simplicity.

To understand this, I called Andrew Anker, who is the chairman of Ebates, an eight-year-old service that offers shopping rebates much as Microsoft is doing now. (Mr. Anker is also a senior vice president of Six Apart, the blog software company.)

Mr. Anker didn't want to denigrate the concept of shopping rebates because Ebates is a profitable small business, with revenue in the low double-digit millions.

"This is a nice business, and we're happy with it," he said. "But if Microsoft thinks this is going to change the game, they will have a hard time of it."

The problem has been mass appeal.

"There are definitely value shoppers who love this sort of thing," he said. "But we haven't figured out how to crack the code to make it a mass business."

I asked him whether Microsoft will have the marketing clout to popularize the concept in a way that Ebates never did. No, that's not it, he replied, noting that his company and others had hooked up with big partners to promote their offers.

Ultimately, he said, most customers didn't find it worth the complexity and time needed to earn a few percent off their shopping.

"People don't want to deal with something that says 'Buy from this place and you will get money back from that place,'" he said. "Search is about how fast can I get someplace: Google's 'I'm feeling lucky' button is a great experience."

To my mind, this is the challenge that Microsoft faces. Nearly two-thirds of Internet users have decided that going to Google is the easiest way to find what they want online. (Indeed, comScore said today that Google's share of searches in the United States increased by 1.8 percentage points in April, to 61.6 percent.) It is going to take more than some rebate checks to convince most people that they should change a behavior that is fast, effective and satisfying.

Text 3: America's Shrinking Food Wraps

<http://www.time.com/time/nation/article/0,8599,1818761,00.html>

American supermarkets are epics of excess: it often seems like every item in the store comes in a "Jumbo" size or has "Bonus!" splashed across the label. But is it possible that the amount of food Americans are buying is, in fact... shrinking? Well, yes. Soaring commodity and fuel prices are driving up costs for manufacturers; faced with a choice between raising prices (which consumers would surely notice) or quietly putting fewer ounces in the bag, carton or cup (which they generally don't) manufacturers are choosing the latter. This month, Kellogg's started shipping Apple Jacks, Cocoa Krispies, Corn Pops, Froot Loops and Honey Smacks containing an average of 2.4 fewer ounces per box.

Similar reductions have recently happened or are on the horizon for many other products: Tropicana orange juice containers are shrinking from 96 ounces to 89; Wrigley's is dropping its the 17-stick PlenTPak in favor of the 15-stick Slim Pack; Dial soap bars now weigh half an ounce less, and that's even before they melt in the shower. Containers of

Country Crock spread, Hellmann's mayonnaise and Edy's and Breyer's ice cream have all slimmed down as well (although that may not necessarily be a bad thing).

"People are just more sensitive to changes in price than changes in quantity," says Harvard Business School Professor John Gourville, who studies consumer decision-making. "Most people can tell you how much a box of cereal costs, but they have no clue how much is actually in it." Other segments of the economy have made similar moves to pass on their higher costs to the consumer without raising prices directly. American Airlines announced in May that it would charge \$15 each way for a single checked bag, part of what airlines have dubbed "a la carte" pricing, which — along with the industrywide drive to put price tags on former freebies like soft drinks, meals and headphones — some airline observers say is really an effort to avoid increasing base ticket prices.

Once they're asked about the changes, food manufacturers are quick to explain their own increasing overhead costs — a Kellogg's spokeswoman said reducing the amount of cereal per box was "to offset rising commodity costs for ingredients and energy used to manufacture and distribute these products" — but most are not exactly going out of their way to let consumers know they're getting less for their money. Some claim newly shrunk products are responses to consumers' needs. Tropicana told the New York Daily News earlier this month that its orange juice containers, which also include a newly designed cap and retail for the same price as the previous larger size, were the result of customer complaints. Said spokeswoman Jamie Stein, "We had a lot of spillage with our old products. It's a value-added redesign."

Reducing the size of products as a way of increasing prices is not new. Frito-Lay cut the amount of chips in their bags and Poland Springs reduced its water cooler jugs from 6 to 5 gallons years ago, all while keeping prices the same. Still, says Chris Waldrop, director of the Food Policy Institute at the Consumer Federal of America, "What's going on now is definitely reflective of rising food costs and rising fuel costs." Waldrop says he doesn't blame manufacturers for taking the step to protect their bottom lines, but says the food companies should be honest with their customers about it. "If they're transparent and open, consumers are less willing to think [manufacturers] are trying to pull one over on them," says Waldrop. The changing product sizes are part of the reason the Bureau of Labor Statistics says groceries cost 5.8% more than the same time last year. Price checkers in the department measure more than 2,000 food items to determine overall food

inflation, and when they notice product size changes, they adjust the inflation index accordingly, according to Ephraim Leibtag, an economist with the Economic Research Service of the Department of Agriculture. ‘

When a product amount drops below a benchmark like "1 pound" or "1 gallon" consumers often take note, according to Gourville. But after that, it's much easier for manufacturers to further whittle down amounts. It's all about taking away consumers' ability to compare apples to apples. The best way to compare food products if you're not sure if sizes have changed is to look at the "unit price," which breaks down the cost per ounce or per quart.

Text 4: Coca-Cola's Big Fizzle (Tuesday, Apr. 12, 2005)

<http://www.time.com/time/magazine/article/0,9171,1048370-1,00.html>

Wednesday, July 10, was C day in America. C for Coca-Cola. C for consumers. C for choice. It was the day that a powerful company in Atlanta felt compelled to return to Americans their national drink. When Coca-Cola announced last April that it was changing the taste of the world's most popular soda, it failed to foresee the sheer frustration and fury that the news would create. From Bangor to Burbank, from Detroit to Dallas, tens of thousands of Coke lovers rose up as one to revile the suddenly sweeter taste of their favorite beverage and demand old Coke back.

Stung and swallowing hard, Coca-Cola reclaimed its birthright last week. In the most spectacular about-face since Ford walked away from its ill-fated Edsel in 1959, the company bowed to public pressure. It declared that old Coke would be restored to groceries, fountains and vending machines within a few weeks. At the same time, the firm said it intended to have its soda and drink it too. Old Coke will return as Coca-Cola Classic. The new Coke that ignited the outrage will remain the flagship brand.

Coke's decision brought forth a joyous response from soft-drink fans across the U.S. Said Karen Wilson, 28, who last June led a rally to protest the new Coke in San Francisco's Union Square: "At first I was numb. Then I was shocked. Then I started to yell and scream and run up and down." Archrival Pepsi professed to be just as delighted. Crowded Roger Enrico, president of Pepsi-Cola USA, about new Coke: "Clearly this is the Edsel of the '80s. This was a terrible mistake. Coke's got a lemon on its hands, and now they're

trying to make lemonade." On Wall Street, though, Coke jumped \$2.37 a share on the announcement, while PepsiCo stock sagged 75¢. In Washington, Democratic Senator David Pryor of Arkansas, an admitted Cokeaholic, expressed his jubilation on the Senate floor. In a speech between a debate over disinvestment in South Africa and action on the Safe Drinking Water Act, Pryor called Coke's capitulation "a very meaningful moment in the history of America. It shows that some national institutions cannot be changed."

Indeed they cannot. As Coke discovered to its sorrow, fiddling with the formula for the 99-year-old beverage was an affront to patriotic pride and perhaps more. "Some people felt that a sacred symbol had been tampered with," said Robert Antonio, a University of Kansas sociologist. Glenwood Davis, marketing manager for Coca-Cola Bottling in Roanoke, Va., said that he received a letter from a woman who said, "There are only two things in my life: God and Coca-Cola. Now you have taken one of those things away from me."

Even after the decision to bring back Classic Coke, company officials were still not quite sure what had hit them. "We did not understand the deep emotions of so many of our customers for Coca-Cola," said President Donald R. Keough. "It is not only a function of culture or upbringing or inherited brand loyalty. It is a wonderful American mystery. A lovely American enigma. And you cannot measure it any more than you can measure love, pride or patriotism."

Everything looked different on April 23, when Coca-Cola Chairman Roberto Goizueta introduced the new Coke, which the firm called "the most significant soft-drink development in the company's history." Gushed Goizueta at the time: "The best has been made even better."

Coke's change was immediately greeted by angry protest. For three straight months, Coca-Cola headquarters received some 1,500 phone calls daily, as well as a barrage of angry letters. Wrote one correspondent: "Changing Coke is like God making the grass purple or putting toes on our ears or teeth on our knees." Among the most common complaints: new Coke was dull and watery and tasted distressingly like Pepsi.

As a chastened Keough admitted last week, "The passion for original Coke was something that just flat caught us by surprise. The simple fact is that all of the time and money and skill poured into consumer research on the new Coca-Cola could not measure

or reveal the depth and emotional attachment to the original Coca-Cola felt by so many people."

Keough denied a widely held belief that the company had brought out new Coke as part of a deliberate, Machiavellian plot to create support for the older product. Said he: "Some critics will say Coca-Cola has made a marketing mistake. Some cynics say that we planned the whole thing. The truth is, we're not that dumb, and we're not that smart."

Why, at a time when the world is trying to cope with international terrorism, superpower tensions and starvation in Africa, has so much furor erupted over what is, after all, merely a change in a soft-drink formula?

In many ways, the Great Coke Debate revealed something about the current state of the American psyche. In a world of ceaseless change, people cling desperately to the known and the given. The old Latin Mass is gone, the phone company has been broken up, Walter Cronkite is no longer on the evening news. Throughout those changes, Coke was always there, a misty memory from childhood, a rock of ages. "Certain things in our psychological environment have to stay constant because we're in such a changing world," says Dr. Bert Pepper (no relation to the soft drink), a New York City psychiatrist. "Each of us has our favorite object of constancy. Many Americans have picked Coke." Adds Pepper: "People felt outraged and ripped off because there was an implicit and explicit contract between the Coke drinker and the company. There was unilateral abrogation of that contract when the company changed the formula."

It is, of course, possible that Coke can turn its near disaster into a marketing coup. The company now has two Cokes to compete with Pepsi-Cola, as an industry watcher pointed out – one that tastes like Coke and one that tastes like Pepsi. And since the soft-drink maker will still be selling new Coke, none of the millions of dollars spent to launch that product has been wasted. If anything, the furor created by the flavor change has made Coke more of a household word than ever.

Appendix 2

Questions and Their Answers in Test 2

1. Which of the following is NOT mentioned in the passage? C
A. Saving Private Ryan
B. The Rocky Mountain News
C. Major Sheeler Beck
D. United States Marine Corps
2. Which of the following contradicts statements in the passage? D
A. All the articles in Sheeler's book are real stories.
B. Major Beck felt it was him who brings the pain to the families.
C. There are specific instructions about how to perform casualty notifications.
D. The content of the book is organized coherently as a series of newspaper articles.
3. Which of the following is not one of Major Beck's personal routines?
A
A. helping to organize dead marines' funerals
B. learning dead marines' names and nicknames
C. touching dead marines' toys
D. reading the letters dead marines wrote home
4. Why are two people required for each casualty notification visit? C
A. Two people can alternate the responsibility of delivering the bad news.
B. Having two representatives visit the families shows more respect than having only one.
C. Family members of dead soldiers may react violently to people who deliver the bad news.
D. When two people are present they can act as more reliable witnesses of the families' reactions.
5. The editor's comment "you have caring in common" is quoted to B
A. emphasize Sheeler's sympathy with families who lose their beloved ones.
B. criticize an inappropriate statement in Sheeler's book.
C. show Major Beck's dedication to his job.
D. provide the perspective of Sheeler's colleagues.
6. Why did the author think Sheeler's book "captures the universal as much as it does the personal"? D
A. The book depicts the grieving process that families endure.
B. The book allows readers to feel the pain experienced by individual families facing loss.
C. The book demonstrates that elements of grief occur in all cultures.
D. In addition to differences, it conveys the similarities of families' responses to loss.

7. What is the author's purpose in mentioning Steven Spielberg's movie at the beginning of this passage? C
- A. To attract readers' attention to a particular scene in the movie.
 - B. To demonstrate the dramatic power of this scene.
 - C. To foreshadow the power of stories in Sheeler's book.
 - D. To evoke readers' sympathy for families who lose their beloved ones.
8. From information in the passage one can infer that B
- A. The author Sheeler is a very systematic thinker.
 - B. It may be difficult to find volunteers to do Major Beck's job.
 - C. The war in Iraq has produced more casualties than expected.
 - D. Major Beck believes that entering the war was a mistake.
9. The skills needed in Major Beck's job are most similar to those of a D
- A. military leader
 - B. teacher
 - C. psychotherapist
 - D. priest
10. Which website is NOT mentioned in the passage? D
- A. eBay
 - B. Amazon
 - C. Yelp
 - D. Yahoo
11. Which company mentioned in the passage does NOT provide reward programs to their customers? C
- A. Upromise
 - B. Ebates
 - C. comScore
 - D. AllAdvantage
12. Which of the following is NOT mentioned as a feature of Google's search services? B
- A. It is easy to use.
 - B. It is free to use.
 - C. It is convenient to use the button "I am feeling lucky".
 - D. It provides fast search results.
13. In the statement "if Microsoft thinks this is going to change the game they will have a hard time of it", "game" refers to a competition for consumers in which market? B
- A. online shopping
 - B. Internet search services
 - C. rebate services

D. social networking

14. Which of the following contradicts the author's opinions in the passage? B
- A. Microsoft is known for taking others' good ideas for their own use.
 - B. Microsoft has enough money to pay people to use its online searching service.
 - C. Mr. Anker thinks providing rebates for online shopping is a good business model to profit.
 - D. People could somehow be influenced by others' behaviors on social networks.
15. The author of the article probably believes that D
- A. searching for online products accounts for a large amount of the search market.
 - B. websites providing services for online products shopping can not survive.
 - C. people want to get shopping discounts while doing online searching.
 - D. appealing to highly specialized segments of the market is a poor strategy for search service providers.
16. Which of the following can NOT be inferred from the passage C
- A. Mass appeal is not a problem for Google.
 - B. Changing users' Internet search behavior demands both technical and psychological conviction.
 - C. The solution to appealing more online shoppers is on the way.
 - D. People are more motivated to participate in a social network if they see others online.
17. What is the author's tone toward Microsoft's rebate-offering program in the passage?
D
- A. indifferent
 - B. unhappy
 - C. worried
 - D. sarcastic
18. If you were Microsoft's marketing manager, which of the following strategies you think can best boost Microsoft's Internet search market? D
- A. simplifying the procedure for customers to redeem their shopping rebates
 - B. increasing the amount of cash back for customers who use Microsoft's Internet search services
 - C. making the search experience as easy as Google's
 - D. partnering with companies that have shares in Internet search market such as Yahoo
19. Which of the following is NOT mentioned as food products in the passage? C
- A. Apple Jacks
 - B. Froot Loops
 - C. Wrigley
 - D. Honey Smacks
20. How much has the price of groceries increased since last year? C
- A. 5.5%

- B. 6.5%
- C. 5.8%
- D. 6.8%

21. According to the Tropicana spokeswoman, the purpose of redesigning the new orange juice container was to D

- A. contain more juice.
- B. cover the increasing overhead costs.
- C. make it look more attractive to customers.
- D. make it more usable for customers.

22. Why did the author think smaller containers of ice cream may not be a bad thing? B

- A. Because consumers can more easily eat ice cream on the go.
- B. Because consumers' intake of fat and calories may be reduced.
- C. Because it costs customers less per package.
- D. Because it will help the manufacturers to stay in business.

23. The author of the passage probably agrees that D

- A. Reducing package sizes is a new strategy for the food industry.
- B. There is no easy way for customers to find out the change of a price for a product.
- C. Food manufactures are not willing to explain their changed sizes in their products.
- D. Tropicana is trying to conceal that they reduced the size of the package to lower their costs.

24. Which of the following contradicts the author's opinions in the passage? C

- A. Most food manufacturers are not willing to admit the reduced amounts in their products.
- B. Increased prices in commodity contribute to the shrinking size of the food products.
- C. Customers can not always tell changes in a product's size when the price of the product remains the same.
- D. Reducing the size of products is an indirect way to increase products' prices.

25. What is the author's purpose in quoting the statement "food companies should be honest with their customers"? B

- A. To warn the food companies not to cheat on their customers
- B. To show that the the food companies' strategy for changing packaging is not necessarily wrong
- C. To show his/her own position on food companies' changes in their product sizes
- D. To show customers' discontentment with not telling them the truth

26. From the information in the passage one can infer that B

- A. the situation in the food industry is well understood by customers
- B. reducing products' sizes can't be a long-term strategy used by food companies to keep their prices stable
- C. food companies are quick to respond to customers' complaints

D. the food industry is most strongly influenced by the rising fuel costs

27. The strategy of reducing sizes of products instead of raising prices is most similar to

C

A. working fewer hours for the same salary

B. commuting that takes longer for the same bus fee

C. offering fewer weekday air minutes for the same monthly cell phone fees

D. providing longer tutorial hours for the same tuition fees

Appendix 3

Questions and Their Answers in Test 3

1. Which of the following does NOT stand for “C day in America” mentioned in the passage? C
A. Coca-Cola
B. consumers
C. classic
D. choice
2. Why did the Coca-Cola company restore its sales of old Coke? B
A. It was planned to draw Coke lovers’ attention to their new Coke products.
B. It yielded to the public pressure to bring back the old taste of Coke.
C. It was deliberately used to boost the sales of Coca-Cola.
D. It was demanded by a democratic senator in Washington.
3. Which of the following are not mentioned in the passage? C
A. Soft-drink lovers were shocked to hear the news of a change in the taste of Coca-Cola.
B. Coke’s stock rose when the old Coke was restored.
C. Coca-Cola’s formula is not known to their Coke fans.
D. People prefer certain favorite objects to stay the same.
4. In the statement “Some cynics say that we planned the whole thing. The truth is, we’re not that dumb, and we’re not that smart”, what does “smart” mean? B
A. Coca-Cola’s officials are not so smart as to take advantage of the public to make profits.
B. Coca-Cola’s officials are not as smart to foresee public’s fury to Coke’s flavor change.
C. Coca-Cola’s officials are not smart enough to make an appropriate decision for promoting their products.
D. Coca-Cola’s officials are not smart enough to produce a better flavor for their Coke customers.
5. Which of the following contradicts the information provided in the passage? B
A. The fury about the flavor change revealed people’s deep emotional connections with Coca-Cola.
B. Coca-Cola officials never understood why people were so angry with the change of Coke’s taste.
C. Coca-Cola is more competitive than before with its two Cokes to compete with Pepsi-Cola.
D. The million-dollar project invested in consumer research on new Coke predicted people’s taste preferences.
6. The author of the passage probably agrees that C

- A. Many Americans feel disloyal if they switch to the new Coke.
 - B. The contract between the Coke drinker and the company is explicitly agreed to.
 - C. Protests caused by the flavor change were an unexpected gain because the Coca-Cola company learned Americans' deep attachment to their product.
 - D. Coca-Cola is also quietly changing to adapt to a world full of changes.
7. What can you infer from author's comment "The company now has two Cokes to compete with Pepsi-Cola, as an industry watcher pointed out--one that tastes like Coke and one that tastes like Pepsi." C
- A. Coca-Cola may lose its distinction from Pepsi-Cola.
 - B. Coca-Cola may have to spend more money in marketing its new Coke.
 - C. Coca-Cola may be able to attract Pepsi-Cola drinkers to become their loyal customers.
 - D. Coca-Cola may unite with Pepsi-Cola to share the soft-drink market.
8. Information in the passage implies that B
- A. people are less concerned about international terrorism than Coca-Cola's flavor change.
 - B. for a company, staying unchanged in a changing world may help build a lasting relationship with their customers.
 - C. Coca-Cola's marketing blunder showed the company it cannot afford to make mistakes
 - D. People are almost always unwilling to try new drinks.
9. What would be Pepsi officials' most plausible feelings about the entire episode of Coke's flavor change? D
- A. Gloating
 - B. Depressed
 - C. Happy
 - D. Worried

Appendix 4

Questionnaire

CIRCLE THE LETTER THAT ANSWERS EACH QUESTION

1. What is your gender?

- a) Male
- b) Female

2. How many credit hours have you completed?

- a) 0-30 credit hours
- b) 31-60 credit hours
- c) 61-90 credit hours
- d) 91-120 credit hours (a bachelor's degree)
- e) more than 120 credit hours

3. What is your major? Please specify

4. During the reading period, how many articles did you restudy?

- a) None
- b) 1
- c) 2
- d) 3
- e) All 4

5. In general, what kinds of information did you restudy?(you can select multiple answers)

- [a] The part(s) in which I am most interested
- [b] The part(s) that I think are difficult
- [c] The part(s) about which I was confused
- [d] The part(s) I thought I might not recall for the tests
- [e] I just reread the whole article again
- [f] None of the above describes what I restudied.
- [g] I did not restudy any part(s) of the 4 readings

6. What part(s) of assigned readings do you usually choose to restudy in your courses?(you can select multiple answers)

- [a] The part(s) that I am the most interested in
- [b] The part(s) that I think very difficult

- [c] The part(s) that I feel very confused
- [d] The part(s) that I feel I cannot recall in the future test
- [e] I only do reread occasionally while studying
- [f] I never reread a reading unless I am asked to do so

7. Restudying helps me learn better.

- a) Strongly agree
- b) Agree
- c) Neutral
- d) Disagree
- e) Strongly disagree

8. Summarizing an article helps me evaluate how well I understand it.

- a) Strongly agree
- b) Agree
- c) Neutral
- d) Disagree
- e) Strongly disagree
- f) Not applicable

9. Summarizing an article helps me plan how to restudy so I can learn more.

- a) Strongly agree
- b) Agree
- c) Neutral
- d) Disagree
- e) Strongly disagree
- f) Not applicable

10. In the future, how likely are you to use restudying as a method for judging how well you understand what you study?

- a) Very likely
- b) likely
- c) Neutral
- d) Maybe not likely
- e) Definitely not likely
- f) Not applicable

11. In the future, how likely are you to use summarization as a method for judging how well you understand what you study?

- a) Very likely
- b) likely
- c) Neutral
- d) Maybe not likely
- e) Definitely not likely
- f) Not applicable