Elaborative retrieval by concept mapping: Is it an effective study tactic?

by

Xin Fang

B.Sc., Guangzhou University of Chinese Medicine, 2008

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Arts

in the
Educational Psychology Program
Faculty of Education

SIMON FRASER UNIVERSITY

Summer 2014

Approval

Name:	**Xin Fang**
Degree:	**Master of Arts**
Title:	***Elaborative retrieval by concept mapping: Is it an effective study tactic?***
Examining Committee:	**Chair:** Dr. Kevin O'Neill Associate Professor

Dr. John Nesbit
Senior Supervisor
Professor

Dr. Phil Winne
Supervisor
Professor

Dr. Elina Birmingham
Internal/External Examiner
Assistant Professor

Date Defended/Approved: August 01, 2014

Partial Copyright Licence

Ethics Statement

SFU

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, collaborator or research assistant 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.

Simon Fraser University Library
Burnaby, British Columbia, Canada

Abstract

The purpose of this experimental research was to investigate studying strategies for learning transferable knowledge. The research involved a comparison of three study strategies (elaborative concept mapping, retrieval by writing, retrieval by concept mapping) with respect to their effects on recall and inference performance. A sample of 120 participants was randomly assigned to four conditions: concept mapping with no retrieval practice, retrieval practice using handwriting, retrieval practice using handwriting with concept map training, and retrieval practice using concept mapping. A posttest consisting of 14 short-answer recall questions and 5 short-answer inference questions was conducted. Among the four conditions there was no statistically significant difference found in the measured outcomes. In accordance with prior research, this study suggests that retrieval practice by concept mapping is an alternative study strategy that has approximately the same effectiveness as retrieval practice by writing. Possible reasons for discrepancies between other aspects of these results and prior research are discussed.

Keywords: elaboration; concept mapping; retrieval practice; learning; study strategies

To my parents

Acknowledgements

I would like to express my special appreciation and thanks to my supervisor Dr. John C. Nesbit, you have been a tremendous mentor for me. I would like to thank you for your patience, motivation, enthusiasm, and immense knowledge. It was your patient guidance and encourage that helped me in all the time of research and writing of this thesis. Besides my supervisor, I would like to thank Dr. Phil H Winne for your careful revisions, insightful comments and hard questions.

I further express my appreciation to Farimah Salimi and Zahia Marzouk. Thank you for sharing your knowledge and experience as well for the support while I was conducting the research.

A special thanks to my family: my parents Jiansheng Fang and Min Chen. Words cannot express how grateful I am to my mother and father for all of the sacrifices that you've made on my behalf. Love you forever.

Last but not least, I also want to thank my friends Xue Peng and Shanshan Zhang. Thank you for being here with me all the way and always backing me up.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1. Rationale

Concept mapping has been found effective for students in obtaining learning outcomes (Winn, 1991; Slotte & Lonka,1999; Novak, & Cañas, 2008; Haugwitz, Nesbit, & Sandmann, 2010). The effectiveness of concept mapping as a study strategy is attributed to elaborative cognitive processing which occurs when students construct concept maps in the presence of the materials they are learning. Elaboration has been supported by research that showed it could improve one's memory for what is being learned (Pressley & Levin, 1987; Rohwer, 1973). In theory, cognitive elaboration promotes meaningful learning by having students mindfully control and transfer what is learned.

However, Karpicke and Blunt (2011a) found that retrieval practice produced greater learning than elaborative studying with concept mapping. Retrieval practice has been found to be a powerful means for helping students to retain knowledge in long term memory (Gates 1917; Karpicke & Roediger, 2007; McDaniel, Howard & Einstein, 2009). In addition, other research has found that the effectiveness of retrieval practice is not likely due to elaborative encoding induced during retrieval-based learning activity (Karpicke & Smith, 2012).

Could using concept mapping as a medium for retrieving studied information benefit learning? Much research has focused on comparing the differences between groups using retrieval practices and repeatedly studying the material without retrieval (Karpicke & Roediger, 2007; Karpicke & Roediger, 2008; McDaniel, Howard & Einstein, 2009). Fewer studies have considered the distinction between retrieval and elaboration strategies, and only one investigated whether combining retrieval practice with concept mapping augments the positive influence these practices have on learning. Blunt and Karpicke (2014) investigated whether concept mapping could be used as a retrieval-based learning activity. The research presented in this thesis has some similarity with the work of Blunt and Karpicke (2014) study. In both studies, retrieval practice using concept mapping was compared with the more conventional type of retrieval practice which uses text (handwritten or typed) as the retrieval medium.

1.2. The Experiment

A total of 120 participants enrolled in Simon Fraser University took part in this research. They were randomly assigned to one of the four conditions: concept mapping with no retrieval practice, retrieval practice using handwriting but with concept map training, retrieval practice using concept mapping, retrieval practice using handwriting. After they studied the presented reading materials by following the assigned procedures, the participants were asked to watch YouTube videos for 22 minutes and then take a posttest about the previous reading materials. Two types of questions: recall and inference questions were included in the posttest to assess different depths of conceptual knowledge. Scores on two types of questions were collected in this study.

1.3. Research Questions

Two research questions were addressed in this research:

(1) As evaluated by recall of and reasoning about studied materials, is there a difference in the instructional effectiveness of conventional retrieval practice (using writing as the expressive medium), concept mapping, and retrieval practice using concept mapping as the expressive medium;

(2) If retrieval practice by concept mapping is more effective than conventional retrieval practice is this due to prior training in concept mapping or to the explicit use of concept mapping while studying?

1.4. Findings

No differences were statistically detected among the four groups in the recall questions and the inference questions. Although there was no significant difference, an examination of the sample means showed that participants who studied by retrieval practice using writing outperformed other groups in the recall posttest. While some aspects of these results are consistent with prior research (Karpicke & Blunt, 2011a; Blunt & Karpicke, 2014) other aspects are not consistent. Several plausible reasons for the discrepancy are discussed in Chapter 5.

Chapter 2. Review of Theory and Prior Research

The research of this thesis investigated studying strategies to improve learning of transferable knowledge. It involved a comparison of three study strategies (elaborative concept mapping, retrieval by writing, retrieval by concept mapping) with respect to their effects on recall and inference performance. It built on prior research and theory about elaborative studying, concept mapping and retrieval practice. In this chapter, a review of theories and prior research on elaborative study, concept mapping and retrieval practice is presented.

2.1. Elaborative studying

The effectiveness of cognitive elaboration as a study strategy has been supported by research that showed it could improve one's memory for what is being learned (Pressley & Levin, 1987; Rohwer, 1973). Elaboration allows students to mindfully control and transfer what is learned which can promote meaningful learning. The idea of meaningful learning was first introduced by Ausubel (1963; 1968) and advocated by other researchers (Kalyuga, 2009; Mayer, 1979, 1989; Novak, 2010). Meaningful learning occurs when the learner connects the new information to knowledge she has already acquired. There are three requirements for meaningful learning: relevant prior knowledge, meaningful materials, learners choosing to learn meaningfully.

Kalyuga (2009) stated that knowledge elaboration processes were essential for meaningful learning as it allowed learners to use prior knowledge to expand and construct new materials. Mayer (1980) summarized three theories which predict a beneficial effect from elaboration strategies: general motivation theory, attention theory and assimilation theory. General motivation theory states that elaboration strategies can help to increase a learner's interest in the materials leading her to encode information deeply and in more detail. Attention theory states that elaboration serves to draw a learner's attention to the information thereby helping to encode it more deeply. Assimilation theory states that elaboration strategies directly facilitate meaningful learning. Elaboration strategies including meaning-enhancing additions or self-generated materials have been found to enhance students' memory for not only simple factual materials but also complex texts. Mayer (1980) found that assimilation theory was the theory most consistent with results from a series of five experiments investigating elaboration learning strategies.

In addition to elaborations based on simple associations between two words, comparative and integrative elaborations are two common types of elaboration strategies that are used in real-world materials. Comparative elaboration happens when the learner explains the relationship between two concepts in the material. Integrative elaboration happens when the learner explains the relationship between a concept in the new materials and concepts in his or her prior knowledge.

Research has found that studying strategies which link prior knowledge to new information such as verbal elaboration ("say why each fact is true"), imagery elaboration ("create a mental picture") and keyword generation ("think of keywords that relate to this new word") can enhance learning performance (Levin, 1976). Pressley, Wood, et al. (1992) claimed that answering questions about new content is an effective elaboration strategy for learners because when they answer "why" questions they are prompted to relate the new learning content with their prior knowledge (Martin & Pressley, 1991; Willoughby, Waller, Wood, & MacKinnon, 1993; Woloshyn, Pressley, & Schneider, 1992; Wood et al., 1990). Both verbal elaboration and imagery elaboration are consistent with the cognitive theories that explain how knowledge is durably and retrievably storied in long term memory (Levin, 1988).

2.2. Concept mapping

Concept maps are defined as graphical tools for organizing and representing knowledge. They consist of concepts, propositions and linking words. The concepts are usually noun phrases or category names. Linking words are usually verb phrases. A proposition is a meaningful statement represented in a concept map by two concepts connected with linking words or phrases. Dansereau and his colleagues (e.g., O'Donnell, Dansereau, & Hall, 2002), devised a type of concept maps called knowledge maps which label links between concepts with symbols representing the nine relationships shown in Figure 2.2.

Concept maps are often hierarchically ordered such that the most inclusive and general concepts are arranged at the top of the map and more specific concepts are arranged below them. Cross-links are horizontally connected links that help learners to see how a concept in one domain of knowledge is related to a concept in another domain shown on the map. Figure 2.1 shows a concept map drawn by a participant in this study. The participant was required to create a hierarchically-ordered concept map by hand and trained to use Dansereau's knowledge map symbols (Figure 2.2.) to represent relationships between concepts.

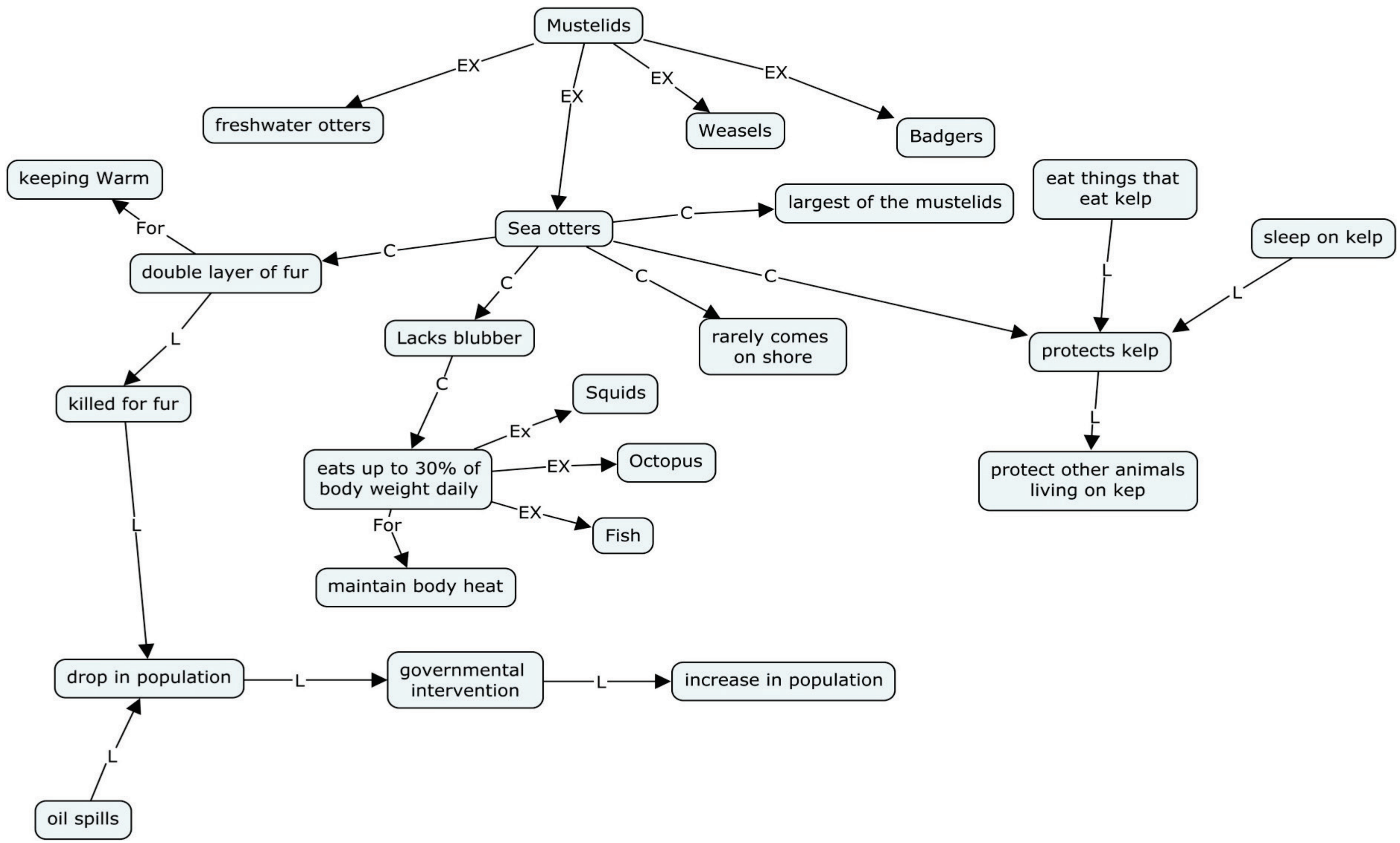


Figure 2.1. **A concept map drawn by a participant in this study (recreated using CmapTools)**

Relationship	Symbol	Example
type	T	planet $\xrightarrow{T}$ gas giant
part	P	comet $\xrightarrow{P}$ tail
characteristic	C	orbit $\xrightarrow{C}$ ellipse
leads to	L	magnetic activity $\xrightarrow{L}$ sunspots
influences	I	mass $\xrightarrow{I}$ force
definition	D	comet $\xrightarrow{D}$ orbiting body with tail
analogy	A	solar system $\xrightarrow{A}$ family
example	EX	dwarf planet $\xrightarrow{EX}$ Pluto
next	N	Mercury $\xrightarrow{N}$ Venus $\xrightarrow{N}$ Earth

Figure 2.2. A list of symbols of nine types of relationships.

Note. Adapted from "Concept maps for learning: Theory, research, and design." by Nesbit, J. C., & Adesope, O. O. (2013). In G. Schraw, M. T. McCrudden, & D. R. Robinson (Eds.), Learning through visual displays (pp. 303–328). Charlotte, NC: Information Age Publishers.

Concept mapping is an active learning task which is thought to trigger cognitive elaboration of new information (Carnot et. al., 2003; Karpicke & Blunt, 2011a). Concept mapping was first developed as a research tool to represent learner's prior knowledge and then deployed as a tool to enhance meaningful learning (Novak, 2010). The idea of concept mapping originally came from Novak's research group as a means of recording what children know about a domain of knowledge before and after instruction. From Ausubel's (1963; 1968) theory of meaningful learning, concept words and propositions given by students would indicate students' prior knowledge and post-instruction knowledge. At first, concept maps only included nodes with concept labels and these were linked in a hierarchical structure without linking words (Moreira, 1977). However, there is a disadvantage in this kind of concept map for people who are unfamiliar with the concepts in the map and therefore do not understand the meaning of the relationships between concepts. Thus, including linking words to express the propositional meanings in a concept map may improve the instructional effectiveness of concept maps. Concept maps can be used as a learning tool either for helping teachers to organize and deliver knowledge or for helping students to represent the ideas in learning materials they are studying. Concept maps can be constructed in various ways. Concept maps can be generated by hand or by a computer program and either by individuals or by groups. Learners can be required to create a concept map while viewing learning materials or create a concept map by retrieving what they have learnt. Concept mapping can be used in a collaborative process. Constructing concept maps by a group can encourage group members' participation in the collaborative process and sharing of information. Novak (2010) claimed that concept maps helped to empower students as learners as students reported that they were "learning how to learn" while experiencing concept mapping.

Research has found evidence for an effect of concept mapping on problem solving and learning outcomes. Slotte and Lonka (1999) found that among applicants who participated in a Finnish medical school entrance examination those who constructed concept maps on scratch paper outperformed other applicants who made verbatim notes or wrote nothing on the scratch paper. Moreover, an experiment by

Haugwitz, Nesbit and Sandmann (2010) showed that groups of students constructing concept maps produced more valid propositions in summaries than groups of students writing summative prose. However, Karpicke and Blunt (2011a) found that students who engaged in retrieval practice learned more than students who engaged in concept mapping. Although they found concept mapping was less effective than retrieval practice, they did not demonstrate that concept mapping was a deficient learning strategy. It may be possible that the way students construct concept maps does not consistently promote elaborative studying (Nesbit & Adesope, 2013). For example, using only superficial cognitive processing, students may be able to extract two noun phrases from a sentence as two nodes and use the verb phrase in the sentence to link the nodes.

The most common explanation for the effectiveness of concept maps as learning tools is that they facilitate meaningful learning and creative thinking. Meaningful learning is deep understanding facilitated by connection of new information to prior knowledge (Ausubel, 1963), in other words, elaborative processing. Meaningful learning is usually conceptualized as the opposite of rote learning which is aimed at producing verbatim recall and avoids deeply understanding new information and making inferences from it. Moreover, concept mapping is sometimes regarded as enhancing the learning capability of learners and increasing recall and transfer of ideas. There are several possible reasons why this may be so. Firstly, concept maps may allow learners to perform more semantic processing in visuospatial working memory and avoid overloading verbal working memory (Winn, 1991; Nesbit & Adesope, 2006, 2013). Secondly, they have simple syntax and may require fewer grammatical decisions than those needed in composing text. This may allow more cognitive resources to be assigned directly to the learning goals (Haugwitz, Nesbit & Sandmann, 2010). Thirdly, compared with other learning strategies (e.g., note-taking, summative prose), concept maps, because of their simple node-link-node structures and hierarchical design, can help the learner to perceive information more immediately and clearly.

2.3. Retrieval practice

Retrieval practice refers to a learning strategy which recalls target information one or more times before assessing learning performance (e.g., examination). Testing and recitation are two important actions that induce retrieval. In 1620, Francis Bacon observed:

> If you read a piece of text through twenty times, you will not learn it by heart so easily as if you read it ten times while attempting to recite from time to time and consulting the text when your memory fails (F. Bacon, 1620/2000, p. 143).

Different from mere repetition, retrieval does not only repeat correct information but must retrieve it when the source is removed from sight. Feedback serves as a function for correcting the response and may be provided during retrieval practice.

Many researchers have shown that learning occurs during testing which involves retrieval practice (Gates, 1917; Bjork, 1975, 1992). A number of experiments showed that testing or repeat-testing practice not only improves learning but also enhances long-term memory due to the act of retrieving information.

How does retrieval affect learning and retention? Various writers have argued that tests requiring recall of materials have greater effects than identification or recognition tests because the former involve greater retrieval effort or depth of processing (Bjork, 1975; Gardiner et al., 1973). Bjork (1975) claimed that depth of retrieval acts like depth of processing during encoding (Craik & Tulving, 1975), and this effortful retrieval may augment the testing effect. Bjork (1992) advocated the notion of creating "desirable difficulties" for learning which can support the testing effect. Bjork's (1992) theory distinguished between storage strength and retrieval strength. Storage strength "reflects the permanence of a memory trace or permanence of learning" and retrieval strength "reflects the momentary accessibility of a memory trace and is similar to the concept of retrieval fluency" (Roediger & Karpicke, 2006a, p.199). In Bjork's model, high retrieval strength (easy retrieval) does not enhance storage strength and low retrieval strength (effortful retrieval) enhances permanence of learning and storage strength. The testing effect is an example of a desirable difficulty because students are required to engage in effortful retrieval in the testing.

Many experimental studies have demonstrated the retrieval practice effect. A classic experiment by Gates (1917) showed the positive effect of retrieval in learning performance. Gates selected a range of grades (Grades 1, 3, 4, 5, 6 and 8) in his study. Two different types of materials (nonsense syllables and brief biographies taken from Who's Who in America) were used in the experiment. Children studied the materials in two phases. In the first phase, children simply read the materials in the assigned time. In the second phase, children were required to recall the information by looking away from materials and reciting it to themselves (covert recitation). Children could glance back at the materials when they sought to refresh their memory. Gates (1917) controlled for amount of time devoted to studying. Children at each age level were divided in different groups to spend 0, 20, 40, 60, 80, or 90% of the learning period doing covert recitation. All groups except first graders with nonsense syllables showed a strong effect of retrieval. With the biography materials, all groups showed an effect. Also, results showed that the strongest retrieval effect occurred when groups spent 60% of the learning period doing covert recitation. The effect decreased when the amount of time spent on retrieval exceeded 60%. Gates concluded that recall attempts during learning as well as a certain amount of time spent on study were a good way to promote learning.

Another experiment originally designed by Tulving (1967) and extended by Karpicke and Roediger (2007) showed that participants undergoing retrieval trials, study-test-study-test (STST) or study-test-test-test (STTT), outperformed participants undergoing only study trials, study-study-study-test (SSST), in long-term retention.

McDaniel et al. (2009) conducted two experiments to evaluate the effect of the read-recite-review (3R) strategy in learning outcomes. The results of two experiments showed that students in the 3R group performed better in initial and delayed free-call tests than students in a reread-only group and a note-taking group. Experiment 1 which used simple prose passages as learning materials revealed no significant differences among three groups in multiple-choice tests and short-answer inference questions.

However, experiment 2 which adopted more complex passages showed that the 3R group performed better on a problem solving test than the rereading group.

2.4. Elaboration and retrieval practice

What cognitive mechanisms might account for the effectiveness of retrieval practice for learning that has been demonstrated in previous research (Gates, 1917; McDaniel, Howard & Einstein, 2009; Roediger & Butler, 2011; Karpicke & Blunt, 2011a)? Some researchers offered the explanation that the retrieval practice process induces elaborative encoding which enhances learning (Carpenter, 2009, 2011; Pyc & Rawson, 2010).

2.4.1. Does retrieval practice induce elaboration?

Karpicke and Smith (2012) presented four experiments that examined whether the mnemonic effect of retrieval practice can be attributed to elaboration. They stated that if the enhancement to long-term retention is due to elaboration, the same effect should occur in the condition of elaboration without retrieval practice. If not, the enhancement is due to retrieval per se. Experiments 1, 2 and 3 compared repeated retrieval with repeated elaboration. Experiments 1 and 2 adopted an imagery-based elaborative study method known as the keyword method and experiment 3 used a verbal elaborative method. In Experiment 4, all subjects learned a list of word pairs which half of the word pairs included different words (mountain – hammer) and half of the word pairs included identical words (castle – castle). When subjects studied a different words pair, they might add additional words or information to link two words and thereby trigger elaborative processing. However, studying an identical-words pair, a subject would have no need to add additional information to link the two identical words and therefore verbal elaboration would be inhibited. If the mnemonic effect of retrieval practice is due to elaboration, the benefit of retrieval practice will be decreased when elaboration is inhibited. On the contrary, if the mnemonic effect of retrieval practice is shown in the groups using retrieval without elaboration, it means that the enhancement of learning can be attributed to retrieval per se.

Experiment 1 and 2 used a similar 3 (learning condition: drop, repeated study, repeated retrieval) x 2 (elaboration condition: elaboration vs. no elaboration) design. In the drop learning condition, a word pair was removed in further study and recall periods once it was correctly recalled. In the repeated study condition, the recalled word pair was kept in further study periods but removed in recall periods. In the repeated retrieval condition, the recalled word pair was kept in further recall periods but removed in study periods. The difference between the two experiments was the time when the keyword mnemonic was provided to the subjects in the elaboration condition. In Experiment 1, the subjects were immediately given an imagery mnemonic after they recalled the definition of the English word. In Experiment 2, an imagery mnemonic was provided during the study trials. The results of the two experiments indicated that there was no significant difference across the elaboration factor. The subjects in the retrieval learning condition performed better than the other two groups whether or not there was elaboration.

The purpose of Experiment 3 was to compare the effect of elaboration to repeated retrieval. There were four conditions: repeated elaboration, drop, repeated study and repeated retrieval. The assumption of this design was that if the effect of retrieval practice was attributed to elaboration, subjects in both the elaboration and repeated retrieval conditions should have the same or similar effect. Significant differences were found among the conditions at the final recall. The subjects in the repeated study and repeated elaboration conditions outperformed the subjects in the drop condition. The subjects in the retrieval condition outperformed the subjects in the elaboration condition. The result of experiment 3 suggested that the mnemonic effect of retrieval practice may not be due to elaboration.

Experiment 4 further examined whether the mnemonic effect of retrieval practice was due to semantic elaboration. All subjects learned a list of word pairs in which half of the word pairs were different words (mountain – hammer) and half of the word pairs were identical words (castle – castle). Elaboration was produced while connecting the target and cue words. By adopting identical-item word pairs, it would reduce or inhibit verbal elaboration. The result showed that the subjects who were given identical word pairs learned much more quickly than the subjects who were given different word pairs. Repeated retrieval enhanced long-term retention both of different word pairs and identical word pairs. For identical word pairs, the final recall test in the repeated retrieval condition was almost perfect.

Overall, the results of Karpicke and Smith (2012) indicated that retrieval practice had a strong mnemonic effect especially for long-term retention, and the effect was due to retrieval per se.

2.4.2. Is retrieval practice more effective than concept mapping with source text in view?

Karpicke and Blunt (2011a) conducted two experiments to compare the effectiveness of retrieval practice and elaborative studying with concept mapping for producing meaningful learning. In Experiment 1, students were arranged in four conditions: study-once, repeated study, elaborative concept mapping and retrieval practice. In the study-once condition, participants studied the text in a single study period. In the repeated study period, participants studied the text in four study periods. In the elaborative concept mapping condition, participants studied the text in an initial study period and then created a concept map. Participants in this condition were instructed about the nature of concept maps and were given an example of a concept map for reference before creating their own concept maps. In the retrieval practice condition, participants first studied the text and then were required to retrieve as much of the content in the text they could. The total amount of learning time in retrieval practice and concept mapping conditions was identical. After participants completed the first retrieval practice, they were required to restudy the text and recall it again. All participants took a final short-answer test one week later. The final short-answer test included verbatim questions and inference questions to assess meaningful learning. The result of the final recall test indicated that the students in repeated study, elaborative concept mapping and retrieval practice conditions outperformed the students in study-once condition. Retrieval practice produced better learning than elaborative studying with concept mapping.

In Experiment 2, all participants were required to create a concept map of one text and practiced retrieval of a second text. For the final test, half of students took a short-answer test and half of students required to create concept maps of the two texts. The final short-answer test used in Experiment 1 was similar to the initial retrieval practice which may be beneficial for participants in retrieval conditions. Thus, creating concept maps in the final test to assess students' learning was added in this experiment. If retrieval practice is beneficial for students to construct concepts and retain the knowledge over the long term, students practicing retrieval should outperform students practicing concept mapping in a final test by creating a concept map. The result indicated that 101 out of 120 students (84%) of the participants performed better on the final text after practicing retrieval than practicing elaborative concept mapping in two final test formats. Results revealed that students involved in retrieval practice produced better performance than students involved in elaborative concept mapping in final short-answer test. In final concept mapping test condition, students in retrieval practice condition produced better performance than students who engaged in concept mapping condition during learning. This suggested that students engaged in concept mapping in learning process did not have advantage when the final test involved concept mapping activity. The result supported retrieval practice as an effective learning strategy.

The findings of Karpicke and Blunt (2011a) strongly support the claim that retrieval practice is a powerful way to improve learning. However, they do not provide unequivocal evidence the concept mapping is an inferior study strategy. Students in the concept mapping group did not receive comprehensive training for how to create a concept map. Students who did not have prior knowledge of concept mapping may incur extra cognitive load during the learning process and may create a concept map in an inefficient way which did not involve much elaboration (Nesbit & Adesope, 2013). More importantly, this research treated concept mapping as an invariant study strategy rather than a mode of verbal expression that, like writing sentences, can be readily combined with various types of study tactics. Students can use retrieval-based concept mapping if they construct concept maps in the absence of the source materials. My research was premised on the idea that retrieval-based concept mapping needs further investigation.

2.4.3. Concept mapping as a medium for retrieval practise

Karpicke and Blunt (2011a) found that retrieval practice produced more learning than concept mapping used as an elaborative study activity. Two experiments conducted by Blunt and Karpicke (2014) further examined the effectiveness of concept mapping used as a retrieval-based practice activity. Some prior research and theory suggested that concept mapping might serve as an effective retrieval-based practice activity. Constructing a concept map requires the student to identify main concepts of materials and link the concepts in an organizational structure (Hay, Kinchin, & Lygo-Baker, 2008; Stewart, Van Kirk, & Rowell, 1979). In addition, students' own prior knowledge is involved in to help them creating concept map (Novak, 1976). Alternatively, concept mapping might not serve as an effective retrieval-based practice activity as it could require additional cognitive load during the process of retrieval. In addition, while students practice retrieval in text format, they tend to recall in serial order. However, practicing retrieval in concept map format might disrupt students' default strategies which could reduce the effect of retrieval practice. Experiment 1 compared the effectiveness of retrieval practice in two formats (paragraph format and concept map format). Each student was required to practice retrieval in paragraph format with one text and practice retrieval in concept map format with another text. In the learning process, students read one text for 5 mins and then recalled it for 10 mins. Then, they reread the text for 5 mins and recalled it again for 10 mins. Students took a final short-answer test (including two types of questions: verbatim and inference) one week later. The result indicated that although students recalled more ideas in paragraph format than in concept map format in the initial performance, performance on the final test was approximately equivalent in both formats. Two repeated study conditions with the text present (paragraph format and concept map format) were added in Experiment 2 to further investigate the effectiveness of retrieval practice. The total amount of learning time was identical in the four conditions. The difference was that students in the repeated study conditions viewed the texts while they completed the learning activities, whereas students in the retrieval practice conditions completed the learning activities with absence of the texts. The result of the experiment indicated that students in retrieval practice conditions outperformed students in repeated study conditions. In the repeated study conditions, a difference in posttest performance favoring concept mapping was found for the verbatim questions but not for the inference questions. This result agreed with prior evidence supporting the idea that concept mapping implemented as an elaborative studying activity produces more learning than only repeated studying. In the retrieval practice conditions, performance on the final test was approximately equivalent for both formats which was similar to the result found in Experiment 1. Overall, Blunt and Karpicke (2014) demonstrated the effectiveness of concept mapping when it was implemented as a retrieval-based practice activity but did not find any advantage of concept mapping over writing as a medium for engaging in retrieval.

Concept mapping has been found to be an effective learning activity (Winn, 1991; Slotte & Lonka, 1999; Novak & Cañas, 2008; Haugwitz, Nesbit & Sandmann, 2010). Retrieval practice has been found to be a powerful means for helping students to retain knowledge in long term memory (Gate 1917; Karpicke & Roediger, 2007; McDaniel, Howard & Einstein, 2009). Research on concept mapping has mainly focused on comparing the differences between writing summaries and concept maps or comparing two concept-mapping techniques (Haugwitz, Nesbit & Sandmann, 2010; Yin, Ruiz - Primo, Ayala & Shavelson, 2005). Research on the retrieval practice hypotheses has mainly compared conditions such as repeated testing, rereading, note-taking or repeated studying and testing (Karpicke & Roediger, 2007, 2008; Roediger & Karpicke, 2006b; McDaniel, Howard & Einstein, 2009). Only a small number of studies, primarily those reviewed here by Karpicke and Blunt, compared the difference between retrieval and elaboration strategies. Karpicke and Blunt's (2011a) study compared retrieval practice and elaborative studying with concept mapping and concluded that retrieval practice was more effective. The finding pointed out that retrieval practice could be used as a powerful learning tool. Also, the study implicitly suggested a new direction for research in which concept mapping could be used as a medium for retrieval practice. The research by Karpicke and Blunt (2014), which pursued that idea and which has some similarity to the research presented in this thesis, was published while data for this thesis research was being collected.

Chapter 3. Method

3.1. Pilot study

A pilot study was conducted with 2 participants to assess the experimental materials and instruments. The purpose of the pilot study was to find out whether the procedures of the experiment were running smoothly. A few problems were discovered. For instance, the initial design of experiment required the participants to play a word game during a 10-minute interpolated task. The intention was to distract the participants from the learning materials and allow for some forgetting to occur. However, the pilot participants said that the word game was not attractive enough and they were so bored after playing it for 5 minutes that they started trying to review what they learnt during the study phase. In addition, the instructions of experimental procedures given to the pilot participants were not clear and concise. These problems were corrected in the main experiment. For the interpolated task, participants were instructed to watch videos instead of playing the word game. Also, a simple table was used to explain the experimental procedures instead of using paragraphs. The pilot participants were excluded from participating in the main experiment.

3.2. Participants

One hundred and twenty students from Simon Fraser University volunteered to participant in this study. The participants were recruited in two ways. First, some participants were recruited when I attended a large enrollment undergraduate course. After informing students of the research topic and procedure, students who expressed interest were contacted via email and scheduled to participate. Second, flyers which described the research topic and procedure were posted around the campus. Participants who were interested contacted the researcher and were scheduled to participate. All participants signed a consent form at the beginning of the experiment and received $15 after completing the study activities. All participants were randomly assigned into one of four conditions: concept mapping with no retrieval practice, retrieval practice using text, retrieval practice using text with concept map training, and retrieval practice using concept mapping.

3.3. Materials and Instruments

The primary reading material was selected from the reading comprehension section of a test-preparation book for the Test of English as a Foreign Language (TOEFL; Rogers, 2001). The passage, provided in Appendix C, covers a single topic ("Sea otters") and is 275 words in length. The passage was originally used by Karpicke and Blunt (2011a). We modified one sentence to avoid confusion that was noticed during the pilot study.

I used the same posttest questions as Karpicke and Blunt (2011a) but added three more inference questions. Their study summed over all items to create a single outcome score. I added the three additional inference items so that I could create separate recall and inference scores. This provided the opportunity to examine the effects of the study strategies on two different types of learning outcomes.

The paper-and-pencil posttest based on the "Sea Otters" passage consisted of 14 short-answer recall questions and 5 short-answer inference questions. An example of one of the recall questions is:

> What is Sea Otter fur made up of?

An example of one of the inference questions is:

> What are the resulting consequences if Sea Otters only choose one type of foods (e.g., Clams) to eat? Please list reasons.

All of the posttest questions are provided in Appendix D.

3.4. Procedure

The research was conducted in a laboratory in the SFU Faculty of Education. Participants who received concept mapping training viewed the concepts of concept mapping via powerpoint slides and created their concept maps on paper while reading the text (Appendix B). The participants were given a sheet of paper (Appendix A) that provided examples of two main concepts of concept map and symbols for common relationships. Participants who were later required to create concept maps in a study phase kept this handout for reference. Participants followed the instructions presented in powerpoint slides and could ask the researcher questions. Reading materials and posttest questions were presented in printed form. Participants were required to create concept maps or perform retrieval practice on paper.

The total amount of time spent studying the text passage ("Sea otters") was identical in the four conditions. All participants studied the text passage in a first study phase lasting 11 minutes and second study phase lasting 10 minutes. The second study phase was one minute shorter because pilot data indicated that participants took up to approximately 3 minutes to read the text passage the first time and up to approximately 2 minutes to read it a second time.

The group who studied by concept mapping with no retrieval practice (CT) first received concept map training. In the first study phase the participants were given 3 minutes to read the text passage and then were given 8 minutes to make a concept map while viewing the text passage. In the second study phase, the first concept map was taken away and participants were given 10 minutes to make a second map while viewing the text passage.

The group who studied by retrieval practice using writing (RW) were first given 3 minutes to read the text passage, and then the text passage was taken away and they were given 8 minutes to recall by writing as much of the information from the text passage as they could, in any order. In the second study phase their prior writing was taken away and participants were given 2 minutes to read the text passage again before being given another 8 minutes to recall by writing.

The group who studied by retrieval practice using writing with concept map training (RWT) first received concept map training and then performed the two study phases in the same manner as the RW group.

The group who studied by retrieval practice using concept mapping (RCT) received the same treatment at the RWT group except they used concept mapping instead of writing as the medium for recall.

After completion of the learning activities, participants were instructed to watch a YouTube playlist of movie trailers which lasted 22 minutes. This interpolated task was done to induce a moderate degree of forgetting of the studied information.

Finally, participants were instructed to complete the posttest. Participants were told that there was no time-limit for the posttest. Overall, it took approximately 60 minutes for a participant to go through the whole procedure.

Chapter 4. Result

4.1. Overview of the data collected

The paper-and-pencil posttest referring to the passage “Sea Otters” consisted of 14 recall short-answer questions and 5 inference short-answer questions. To make the scoring process more precise and consistent, the standard answer to each question was divided into key points which assigned different scores depending on their relevance. A scorer unaware of treatment conditions used a scoring rubric to grade the answers.

The posttest was scored using the same method adopted by Karpicke and Blunt (2011a) but slightly modified the scoring standard in one of the recall questions. Also, as mentioned previously, this research used a posttest that included 3 additional items.

There were a total of 30 points assigned for the posttest, 21 points for the 14 recall questions and 9 points for the 5 inference questions. The number of points assigned to each question was different because the expected number of ideas in a complete answer varied. Of the recall questions, the range was from 1 to 7 points per question. Of the inference questions, one was 1-point question and five were 2-points question. For each expected idea, a complete and accurate expression of that idea was scored as 1. A partially stated and accurate expression of the idea was scored as 0.5. An absent or incorrect expression of an idea was scored as 0. The overall posttest score for a participant was calculated by summing the scores for each question. Also, separate summed scores for the recall questions and the inference questions were calculated for each participant.

4.2. Analysis and results

For each of the question types and overall performance, a one-way ANOVA with the four learning conditions comprising the between-subjects factor was conducted.

Before doing analysis, the data had been checked to make sure that the data can be analyzed using one-way ANOVA. To receive a valid result, the data have to match six assumptions:

> Assumption 1: The dependent variable should be measured at the interval or ratio level (i.e., they are continuous).
>
> Assumption 2: The independent variable should consist of two or more categorical, independent groups.
>
> Assumption 3: All observations should be independent. For example, one participant cannot be in more than one group.
>
> Assumption 4: There should be no significant outlier.
>
> Assumption 5: The dependent variable should be approximately normally distributed for each category of the independent variable.
>
> Assumption 6: There needs to be homogeneity of variances.

In this study, the dependent variable was the scores of outcome test and the independent variable included four groups: studied by concept mapping with no retrieval practice (CT), studied by retrieval practice using writing (RW), studied by retrieval practice using writing with concept map training (RWT), and studied by retrieval practice using concept mapping (RCT). Those matched the assumption 1 and 2. In addition, the study collected 120 students from SFU which were randomly assigned in four conditions and no student was in more than one group. The assumption 3 was also matched. To examine whether the data matched the assumption 4, 5 and 6, further explanation will be presented at the following sections.

4.2.1. Outliers and Normality

In this study, the dependent variables were total points from the 14 recall questions, total points from the 5 inference questions and total points for the whole posttest. Every variable in each of the four groups was checked for outliers and normality before doing the analysis. The way to detect the outliers was to convert all of the scores for a variable to standard scores. The sample size of this study was 120 cases which meant that a case was an outlier if its standard score was ± 3.0 or beyond. The range of standard scores of recall questions was from -2.31 to 2.00. The range of standard scores of inference questions was from -2.06 to 1.85. The range of standard scores of total outcome test was from -2.30 to 1.99. No standard score was ± 3.0 or beyond. In conclusion, there was no significant outlier detected in the data.

Normality was also checked to gain a thorough understanding of the shape of the distribution of each variable in four groups. Skewness and Kurtosis are two important factors to examine the shape of normal distribution. Table 4.1 showed that the skewness and kurtosis values were all less than or close to ± 1. This indicated that distributions of these variables were within limit of a normal distribution. The Shapiro Wilk test is another way to examine the shape of normal distribution. A p-value below a pre-determined threshold would indicate that the sample data were drawn from a population that was not normally distributed. As shown on Table 4.1, 2 of the 12 variables failed the test of normality at the $p < .05$ level. However, since 12 tests were conducted, the chance of type-I error was somewhat inflated and a lower threshold would be appropriate. Considering that the p values of those two variables were larger than .01 and their skewness and kurtosis were moderate, I concluded that the distributions of the two variables did not extremely deviate from the normal distribution. It was therefore reasonable to proceed with the three one-way ANOVAs.

Table 4.1. Tests of Normality

	Group	Shapiro Wilks (*p*)	Skewness(SE)	Kurtosis(SE)
Verbatim Question	CT	.344	-.44(.43)	-.25(.83)
	RCT	.599	.21(.43)	-.57(.83)
	RW	.173	-.33(.43)	-.79(.83)
	RWT	.351	.01(.43)	-1.06(.83)
Inference Question	CT	.024	-.18(.43)	-1.20(.83)
	RCT	.012	-.73(.43)	-.28(.83)
	RW	.190	.02(.43)	-.87(.83)
	RWT	.118	-.32(.43)	-.80(.83)
Total	CT	.565	-.25(.43)	-.82(.83)
	RCT	.328	-.30(.43)	-.31(.83)
	RW	.301	-.20(.43)	-.58(.83)
	RWT	.113	-.12(.43)	-1.26(.83)

4.2.2. The analyses of the outcome test

Table 4.2 showed the means and standard deviations for the posttest total score, verbatim questions total score and inference questions total score for the four groups.

A one-way ANOVA indicated that there was no statistically significant difference between groups on the total test score, F (3,116) = .397, p = .755. The largest effect size, d = .24, was between the group who studied by retrieval practice without concept map training (M = 17.22, SD = 5.09) and the group who studied by concept mapping with no retrieval practice (M = 15.82, SD = 6.40).

Table 4.2. Descriptive Statistics

	Group	Mean	SD	N
Total	CT	15.82	6.40	30
(out of 30)	RCT	15.87	5.49	30
	RW	17.22	5.09	30
	RWT	16.68	6.41	30
Recall	CT	11.02	4.19	30
(out of 21)	RCT	11.17	4.17	30
	RW	12.55	3.83	30
	RWT	11.88	4.49	30
Inference	CT	4.80	2.61	30
(out of 9)	RCT	4.70	2.23	30
	RW	4.67	2.04	30
	RWT	4.80	2.41	30

Recall questions

For the recall questions, there was no statistically significant difference between groups as determined by a one-way ANOVA (F (3,116) = .859, p = .465). The largest effect size, d = .38, was between the group who studied by retrieval practice without concept map training (M = 12.55, SD = 3.83) and the group who studied by concept mapping with no retrieval practice (M = 11.02, SD = 4.19).

Inference questions

For the inference questions, there was no statistically significant difference between groups as determined by a one-way ANOVA (F (3,116) = .026, p = .994). The largest effect size, only d = .06, was between the group who studied by retrieval practice with concept map training (M = 4.80, SD = 2.41) and the group who studied by retrieval practice without concept map training (M = 4.67, SD = 2.04).

Chapter 5. Discussion

5.1. How the research design differed from prior research

In some respects, this research can be seen as a replication of experiments by Karpicke and Blunt (2011a, 2014). Therefore, it is important to review how the research design aligned and departed from that prior work before we interpret and compare the results.

A criticism of Karpicke and Blunt's (2011a) research design by Mintzes et al., (2011) focussed on "the relatively brief period of training and practice in concept mapping offered the research participants and the absence of evidence that they actually mastered the technique before the experiment" (p. 453). Participants in Karpicke and Blunt's study were briefly informed about concept mapping as a study strategy and were shown an example of a concept map. Although Mintzes et al., (2011) stated that extensive training is needed for students to master the strategy, Karpicke and Blunt (2011b) rebutted the criticism by citing several articles (Hay, Kinchin, & Lygo-Baker, 2008; Mintzes, Wandersee & Novak, 2001; Quinn, Mintzes & Laws, 2003) which claimed that only a brief period of training is required for students to use and benefit from concept mapping.

Blunt and Karpicke (2014) conducted an experiment which replicated the experiment in Karpicke and Blunt (2011a) but added one more condition: retrieval-based concept mapping similar to the RCT condition in my research. In Karpicke and Blunt (2014), all concept mapping participants received a brief period of training which was as the same as in their previous study.

Although my research provided only 10 minutes of concept map training, the training included three key elements not provided by Karpicke and Blunt. First, the participants were told that concept maps should be hierarchically structured with more inclusive concepts displayed above the more specific concepts to which they were related, and they were provided with an example of a hierarchical relationship (Appendix A). Second, the participants were told that concept maps can use a common set of symbols for describing relationships, and they were provided with a table identifying 9 symbols and showing an example of how each could be used (Appendix A). Third, the participants constructed a practice concept map from a short passage (Appendix B). After they finished the practice map, they were given a model example of a concept map for the same passage which they could compare to their own version.

The total time assigned to studying was different from Karpicke and Blunt's research. Although Karpicke and Blunt allocated a total of 30 minutes for study (including retrieval practice and concept mapping), my research design allocated 21 minutes. Pilot testing suggested that the time allocated for initial reading of the material in the prior research was significantly more than the SFU participants required. Significant 'free' time can threaten the internal validity of a learning experiment, as indicated by pilot participants in the CT condition who chose to use the extra time to engage in practice retrieval. The change in the allocated time allowed potential differences in the learning outcome to be attributed more securely to the intended treatment condition. The total amount of learning time was identical in four conditions as was the total learning time in Karpicke and Blunt's research.

Another difference from Karpicke and Blunt's design was that three inference questions were added to the posttest to provide an opportunity to explore the effects of the four conditions across question types. Last but not least, participants in this study took the final test after 22 minutes distractive activity instead of one week later.

Although some changes have been made in this experiment, I judged that the effectiveness of the treatments was not hampered; and indeed in the case of concept mapping was provided in a manner more compatible with what prior research and theory on concept mapping would recommend. With these variations in mind, it is feasible to compare the result of this experiment with the results of Karpicke and Blunt.

5.2. Discussion of the results

There was no statistically significant difference found in the performance of the outcome test among four conditions in this experiment, although an effect size of $d = .38$ favoring the retrieval practice conditions over the concept mapping conditions on recall questions was in the same direction as Karpicke and Blunt's results (2011a, 2014). Recall questions assess conceptual knowledge represented directly in the text, and inference question require students to connect multiple concepts from the text and fill in information not represented. Both question types are conceptual but assess different depths of conceptual knowledge. The trend of these results hints but of course cannot demonstrate that retrieval practice may have more powerful effects on memory for information directly represented in the source text.

Regarding the comparison of conventional concept mapping with retrieval practice, what can account for the discrepancy between my results and those of Karpicke and Blunt? One explanation is that participants in the concept mapping condition benefitted from changes I introduced in the training and enactment of concept mapping. The opportunity to practice concept mapping in the training phase of my research design may, by itself, have been enough to raise the performance of the concept mapping groups to near the level of the retrieval practice groups. Moreover, participants in concept mapping conditions were required to create concept maps in a hierarchical structure and were encouraged to use Dansereau’s relational symbols to label the links or create their own symbols. These mapping tactics likely required participants to engage in a deeper level of cognitive processing and elaborative studying. A second plausible explanation for the discrepancy is that the superiority of retrieval practice observed by Karpicke and Blunt only manifests over retention intervals longer than the within-session interval in my research design. To test these hypotheses, experiments are needed which compare the different concept mapping training conditions (with or without practice), the types of concept mapping (e.g., with and without hierarchical ordering), and to replicate my research design with longer retention intervals.

Considering that the materials and posttest used in this research were adopted from Karpicke and Blunt (2011) and were only augmented by three inference questions I developed, the scores of participants in the RW conditions were surprisingly low compare with those reported by Karpicke and Blunt. The differing types of incentive may explain this outcome. Karpicke and Blunt (2011) offered the students course credits for participation. In my research, participants received $15 for participation. Participants who received course credits might have been implicitly prompted to have a higher motivation to learn because the research participation was offered in lieu of academic tasks.

Both Blunt and Karpicke (2014) and my study added a retrieval-based concept mapping condition which required participants to create a concept map in the absence of the text. In Blunt and Karpicke (2014), the posttest performance of writing and concept mapping retrieval treatments was statistically equivalent, which is the same result as my research. Contrary to my expectations, there was no difference in learning outcomes between the participants who drew concept maps in the presence of source materials and those who drew them from memory. Given the possibility that retrieval practice manifests more powerfully over longer retention intervals, this is another reason to replicate the present design with the posttest delayed, say, one week after the study session.

5.3. Implications for instruction

There is a significant amount of prior research that has found concept mapping and retrieval practice to be effective learning tactics. If we accept those findings then the results of the present research suggest that the two tactics offer similar levels of efficacy. Both retrieval practice and concept mapping can be adopted in classroom settings and solo studying to facilitate teaching and learning. Perhaps the most natural way of implementing retrieval practice in classrooms is to have students answer questions. For example, at the end of class, teachers can provide a small set of questions that require student to retrieve information from the same day's learning content; or, questions can be posed about information studied in previous classes. Practicing retrieval may not only facilitate students engaging in class activities but also help students to reconstruct and retain knowledge they have previously learned.

There are many applications of concept mapping in education. Firstly, concept map can be used as a tool for support of learning. For example, after reading an assignment or completing some other lessons, concept mapping can be used for note taking or as a study aid to summarize what has been learnt. Moreover, concept maps can be constructed not only by an individual but also by a group as part of collaborative group process. Concept mapping can facilitate the exchange of information in a group and encourage students to participate in the collaborative process.

5.4. Limitations and future work

It is difficult to exclude the possibility that variation in factors such as individual memory capacity and individual prior knowledge might have occluded the effects due to treatment. Unaccounted for variance can often hide the signal of treatment effects in what amounts to statistical noise. If this were the case, a research design that measured such factors would have helped to find the treatment effects.

Another limitation is that collecting data over only a single session may not reflect the longer term development of the learning strategies that were investigated. The participants may not have used the learning strategies before, and, even with the training I provided, they may have felt uncomfortable using unaccustomed learning strategies and that may have hampered their performance.

Furthermore, the participants in this experiment took a final recall test after 22 minutes of distractive activity rather than a week later as in previous research. The effectiveness of tactics for long-term retention may not be fully explored in this study as the time interval between learning and testing is not long enough.

In addition, there may be something about the particular text or science content selected in this study that in some way biased the findings of this research. For example, it might not facilitate participants connecting the presented concepts with their prior knowledge as well as other texts that could have been used. This limitation can only be overcome by examining the effects of concept mapping and retrieval practice over a wide range of study materials and subject areas.

In this study, the format of the posttest was more similar to the retrieval practice condition, and that might have had an effect on the performance. Future studies could investigate whether a match or mismatch between the study medium and testing medium would have an effect on the results. The design of further studies could include two different final test formats (short-answer and concept map).

Also, future research might probe the effectiveness of the learning tactics without controlling study time. Without the control of time, participants who were using unaccustomed learning tactics such as concept mapping would have more time to develop fluency in the tactics which could have an effect on learning. Moreover, participants might concentrate more on the study task without the distraction of time management.

References

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning.* New York: Grune & Stratton.

Ausubel, D. P. (1968). *Educational psychology: A cognitive approach.* New York: Holt,Rinehart & Winston.

Bjork, R. A. (1975). Short-term storage: The ordered output of a central processor. *Cognitive Theory*, *1*, 151-171.

Bjork, R. A. (1992). Interference and memory. In L. R. Squire (Ed.), *Encyclopedia of learning and memory* (pp. 283-288). New York: Macmillan.

Blunt, J. R., & Karpicke, J. D. (2014). Learning With Retrieval-Based Concept Mapping. *Journal of Educational Psychology*, *106*, 849-858. doi:10.1037/a0035934

Carnot, M. J., Feltovich, P., Hoffman, R. R., Feltovich, J., & Novak, J. D. (2003). A summary of literature pertaining to the use of concept mapping techniques and technologies for education and performance support. *Pensacola, FL*.

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(6), 1563.

Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1547.

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268.

Gardiner, F. M., Craik, F. I., & Bleasdale, F. A. (1973). Retrieval difficulty and subsequent recall. *Memory & Cognition*, *1*(3), 213-216.

Gates, A.I. (1917). Recitation as a factor in memorizing. *Archives of Psychology*, *6*(40).

Haugwitz, M., Nesbit, J. C., & Sandmann, A. (2010). Cognitive ability and the instructional efficacy of collaborative concept mapping. *Learning and Individual Differences, 20*(5), 536-543. doi:10.1016/j.lindif.2010.04.004

Hay, D., Kinchin, I., & Lygo-Baker, S. (2008). Making learning visible: the role of concept mapping in higher education. *Studies In Higher Education*, *33*(3), 295-311.

Kalyuga, S. (2009). Knowledge elaboration: A cognitive load perspective. *Learning and Instruction*, *19*(5), 402-410. doi:10.1016/j.learninstruc.2009.02.003

Karpicke, J. D., & Blunt, J. R. (2011a). Retrieval practice produces more learning than elaborate studying with concept mapping. *Science, 331*(6018), 772-775. doi:10.1126/science.1199327

Karpicke, J. D., & Blunt, J. R. (2011b). Response to comment on "retrieval practice produces more learning than elaborative studying with concept mapping". *Science*, *334*(6055), 453-453.

Karpicke, J. D., & Roediger III, H. L. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language, 57*(2), 151-162.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, *319*(5865), 966-968.

Karpicke, J. D., & Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding. *Journal of Memory and Language, 67*(1), 17-29. doi:10.1016/j.jml.2012.02.004

Levin, J. R. (1976). What have we learned about maximizing what children learn? In J.R. Levin & V.L.Allen (Eds.), *Cognitive learning in children: Theories and strategies* (pp. 105-134). New York: Academic Press.

Levin, J. R. (1988). Elaboration-Based learning strategies: Powerful THEORY = powerful application. *Contemporary Educational Psychology*, *13* (3), 191-205

Martin, V. L., & Pressley, M. (1991). Elaborative-interrogation effects depend on the nature of the question. *Journal of Educational Psychology*, *83*(1), 113.

Mayer, R. E. (1979). Can advance organizers influence meaningful learning?.*Review of educational research*, 371-383.

Mayer, R. E. (1980). Elaboration techniques that increase the meaningfulness of technical text: An experimental test of the learning strategy hypothesis. *Journal of Educational Psychology,72*(6), 770-784. doi:10.1037/0022-0663.72.6.770

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

McDaniel, M. A., Howard, D. C., Einstein, G. O. (2009). The read-recitel-review study strategy: effective and portable. *Psychological Science, 20* (4), 516-522

Mintzes, J. J., Canas, A., Coffey, J., Gorman, J., Gurley, L., Hoffman, R., ... & Wandersee, J. H. (2011). Comment on “retrieval practice produces more learning than elaborative studying with concept mapping”. *Science*, *334*(6055), 453-453.

Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (2001). Assessing understanding in biology. *Journal of Biological Education*, *35*(3), 118-124.

Moreira, M. M. (1977). *The Learning Theory of David Ausubel as an Alternative Framework for Organizing the Teaching of Language and Literature.* Ithaca, NY, Cornell University

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*(3), 413-448. doi:10.3102/00346543076003413

Nesbit, J. C., & Adesope, O. O. (2013). Concept maps for learning: Theory, research, and design. In G. Schraw, M. T. McCrudden, & D. R. Robinson (Eds.), *Learning through visual displays* (pp. 303–328). Charlotte, NC: Information Age Publishers

Novak, J. D. (1976). Understanding the learning process and effectiveness of teaching methods in the classroom, laboratory, and field. *Science Education*,*60*(4), 493-512.

Novak, J. D., & Cañas, A. J. (2008). *The theory underlying concept maps and how to construct them* (Technical Report IHMC CmapTools 01-2006, Revised 01-2008, Florida Institute for Human and Machine Cognition). Retrieved August 28, 2008 from http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf.

Novak, J. D. (2010). *Learning, creating, and using knowledge: Concept map as facilitative tools in schools and corporations (2nd ed.).* New York, NY: Routledge.

O'Donnell, A.M., Dansereau, D.F., & Hall, R. H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educational Psychology Review, 14*, 71-86.

Pressley, M., & Levin, J. R. (1987). Elaborative learning strategies for the inefficient learner. *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities*, *2*, 175-212.

Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist, 27*(1), 91-109.

Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, *330*(6002), 335-335.

Quinn, H. J., Mintzes, J. J., & Laws, R. A. (2003). Successive concept mapping. *Journal of College Science Teaching*, *33*(3), 12-16.

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20-27.

Roediger, HL; Karpicke, JD (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspective on Psychological Science, 1* (3), 181-210

Roediger, H. L., & Karpicke, J. D. (2006b). Test-enhanced learning taking memory tests improves long-term retention. Psychological Science, *17*(3), 249-255.

Rohwer, W. D., Jr. (1973). Educational inequality: Race is not the problem. *California Journal of Teacher Education,* 1, 6-19. Reprinted in *Integrated Education*, 1974, 13, 3-9.

Slotte, V., & Lonka, K. (1999). Spontaneous concept maps aiding the understanding of scientific concepts. *International Journal of Science Education*, *21*(5), 515-531.

Stewart, J. (1979). Concept Maps: A Tool for Use in Biology. *Teaching.American Biology Teacher*, *41*(3), 171-75.

Tulving, E. (1967). The effects of presentation and recall of material in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, *6*(2), 175-184.

Willoughby, T., Waller, T., Wood, E., & MacKinnon, G. E. (1993). The effect of prior knowledge on an immediate and delayed associative learning task following elaborative interrogation. *Contemporary Educational Psychology*, *18*(1), 36-46.

Winn, W. (1991). Learning from maps and diagrams. *Educational Psychology Review*, *3*(3), 211-247.

Woloshyn, V. E., Pressley, M., & Schneider, W. (1992). Elaborative-interrogation and prior-knowledge effects on learning of facts. *Journal of Educational Psychology*, *84*(1), 115.

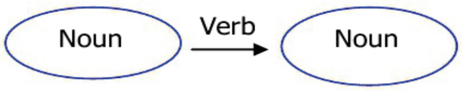
Wood, E., Pressley, M., & Winne, P. H. (1990). Elaborative interrogation effects on children's learning of factual content. *Journal of Educational Psychology*,*82*(4), 741.

Yin, Y., Vanides, J., Ruiz‐Primo, M. A., Ayala, C. C., & Shavelson, R. J. (2005). Comparison of two concept‐mapping techniques: Implications for scoring, interpretation, and use. *Journal of Research in Science Teaching*,*42*(2), 166-184.

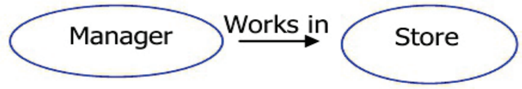
Appendix A.

Handout

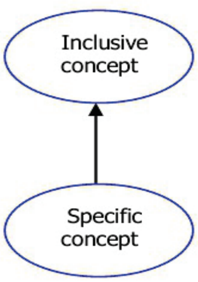
1. Relation



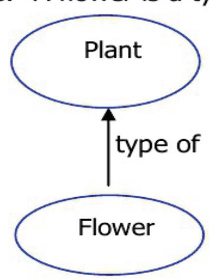
Example: "The manager works in a store."



2. Hierarchical



Example: "A flower is a type of plant."



Symbols for common relationships

Relationship	Symbol	Example
type	T	planet —T→ gas giant
part	P	comet —P→ tail
characteristic	C	orbit —C→ ellipse
leads to	L	magnetic activity —L→ sunspots
influences	I	mass —I→ force
definition	D	comet —D→ orbiting body with tail
analogy	A	solar system —A→ family
example	EX	dwarf planet —EX→ Pluto
next	N	Mercury —N→ Venus —N→ Earth

Appendix B.

Practice text

Climate

Climate is determined by energy transfer from the Sun at and near Earth's surface. This energy transfer is influenced by dynamic processes, such as cloud cover and Earth's rotation, as well as static conditions, such as proximity to mountain ranges and the ocean. Human activities, such as the burning of fossil fuels, also affect the global climate.

Concept Map:

Appendix C.

Studied text passage

Sea Otters

Sea otters dwell in the North Pacific. They are the largest of the mustelids, a group that also includes freshwater otters, weasels, and badgers. They are from 4 to 5 feet long and most weigh from 60 to 85 pounds. Large males may weigh 100 pounds or more.

Unlike most marine mammals, such as seals or dolphins, sea otters lack a layer of blubber, and therefore have to eat up to 30 percent of their body weight a day in clams, crabs, fish, octopus, squids, and other delicacies to maintain body heat. However, their voracious appetites do not create food shortage. This is because each animal is a picky eater and prefers only a few types of food. Thus no single type of food source is exhausted. Sea otters play an important environmental role by protecting forests of seaweed called kelp, which provide shelter and nutrients to many species. Certain sea otters feast on invertebrates, like sea urchins and abalones, that destroy kelp.

Sea otters eat and sleep while floating on their backs, often on masses of kelp. They seldom come on shore. Sea otters keep warm by means of their luxuriant double-layered fur, the densest among animals. The soft outer fur forms a protective cover that keeps the fine underfur dry. One square inch of underfur contains up to one million hairs. Unfortunately, this essential feature almost led to their extinction, as commercial hunters drastically reduced their numbers.

Under government protection, the sea otter population has recovered. However, occasionally unfortunate events have damaged the sea otter population. For example, in 1989, up to 5,000 sea otters perished when the Exxon Valdez spilled oil in Prince William Sound, Alaska.

Appendix D.

Posttest

Questions: Sea Otters

Part I: Factual Questions

1) What is Sea Otter fur made up of?

2) How does a Sea Otter's fur keep it warm?

3) Why do Sea Otters eat 30% of their body weight a day?

4) Why don't Sea Otters encounter food shortages?

5) Sea Otter's play an environment role by protecting something, what is it?

6) Where do Sea Otters spend most of their time?

7) Sea Otters are the largest of which categorical group?

8) How long are Sea Otters?

9) How much do Sea Otters typically weigh?

10) How much can large males weigh?

11) One square inch of a Sea Otter's underfur contains up to how many hairs?

12) Where did the Exxon Valdez oil spill occur?

13) What year did the Exxon Valdez oil spill occur?

14) Seven specific foods that Sea Otters eat where mentioned in the text, name as many as you can.

Part II: Explanation Questions

Answer each of the following questions and then give the explanation or reasons for your answer.

15) What are the resulting consequences in removing Sea Otters from the environment?

16) How is a Sea Otter's most essential survival feature also one of the most dangerous for it?

17) What are the resulting consequences if Sea Otters only choose one type of foods (e.g., Clams) to eat? Please list reasons.

18) What are the resulting consequences for Sea Otters in removing Kelp from the environment?

19) What substance has been found to be toxic to Sea Otters?

Appendix E.

Ethics approval

SFU OFFICE OF RESEARCH ETHICS

Street Address
Simon Fraser University
Discovery 2
Room 230, 8900 Nelson Way
Burnaby, BC Canada V5A 4W9

Mailing Address
8888 University Drive
Discovery 2
Burnaby, BC Canada
V5A 1S6

Director 778.782.6593
Associate Director 778.782.9631
Manager 778.782.3447
FAX 778.782.6785

dore@sfu.ca
http://www.sfu.ca/vp-research/ethics/

Delegated Minimal Risk Approval

Study Number: 2013s0286
Study Title: Elaborative retrieval by concept mapping: Is it an effective study tactic?

Approval Date: 2013 May 15
Principal Investigator: Fang, Xin
SFU Position: Graduate Student

Expiry Date: 2014 May 15
Supervisor: Nesbit, John
Faculty/Department: Education

Co-Investigators: none

Funding Source: none
Grant Title: n/a

Documents Approved in this Application:

- Study Details, version uploaded 2013 May 14
- Consent Form, dated 2013 May 7
- Questions and Answers
- Recruitment Advertisement

I am pleased to inform you that the above referenced study has been approved by the Associate Director, Office of Research Ethics, on behalf of the Research Ethics Board in accordance with University Policy R20.01 (http://www.sfu.ca/policies/research/r20.01.htm). The Board reviews and may amend decisions or subsequent amendments made independently by the Associate Director, Director, Chair or Deputy Chair at its regular monthly meeting.

The approval for this protocol expires on the **Expiry Date**, or the term of your appointment/employment/student registration at SFU, whichever comes first. **An annual renewal form must be completed every year prior to the anniversary date of approval. Failure to submit an annual renewal form will lead to your study being suspended and potentially terminated.** If you receive any grant for this protocol in addition to any funding listed above, please email dore@sfu.ca stating the funding source, the term of approval of the funding source and the title of that funding application if it differs from the title of your ethics application. If you intend to continue your protocol to collect data past the term of approval, you must contact the Office of Research Ethics at dore@sfu.ca and request an extension at least 6 weeks before the expiry date.

The Office of Research Ethics must be notified of any changes in the approved protocol. If you wish to revise your study in any way, please send an email requesting an amendment addressed to dore@sfu.ca. In all email correspondence relating to this application, please reference the application number shown on this letter, which should be included in square brackets at the beginning of the Subject Line; this will ensure that all correspondence is saved to the electronic study file.

Your application has been categorized as "Minimal Risk". "Minimal Risk" occurs when potential participants can reasonably be expected to regard the probability and magnitude of possible harms to be no greater than those encountered by the participant in those aspects of his or her everyday life that relate to the research. Please note that it is the responsibility of the researcher, or the responsibility of the Student Supervisor if the researcher is a graduate student or undergraduate student, to maintain written or other forms of documented consent for a period of 1 year after the research has been completed.

The REB assumes that investigators continuously review new information for findings that indicate a change should be made to the study protocol or consent documents and that such changes will be brought to the attention of the ORE in a timely manner.

If there is an adverse event, the principal investigator must notify the Office of Research Ethics within five (5) days. An Adverse Events Form is available electronically by contacting dore@sfu.ca.

All correspondence with regards to this application will be sent to your SFU email address.

Please notify the Office of Research Ethics at dore@sfu.ca once you have completed the data collection portion of your project so that we can close the file.

This Notification of Status is your official ethics approval documentation for this project. Please keep this document for reference purposes and acknowledge receipt of this Notification of Status by email to dore@sfu.ca and include the study number in square brackets as the first item in the Subject Line.

Best wishes for success in this research.

Sincerely,

[redacted]

Dina Shafey, PhD, MBA
Associate Director
Office of Research Ethics