A STUDY OF SEMI-HIERARCHICAL ORGANIZATION IN THE
CONSTRUCTION OF CONCEPT MAPS USING THE
FRAMEWORK OF COGNITIVE LOAD THEORY

by

DEV THAIN

A thesis submitted to the Faculty of Education
In conformity with the requirements for
the degree of Master of Education

Queen’s University
Kingston, Ontario, Canada
May, 2012

Copyright © Dev Thain, 2012
Abstract

The value of hierarchy as an essential trait of concept maps and a way to enhance recall is explored in this thesis. Undergraduate students ($N = 40$) were randomly assigned to one of two groups and completed an 18-question multiple-choice pre-test about the concept of animal physiology. Then each group studied one of two visual organizers that varied in the level of hierarchy used and finally both groups completed the same multiple-choice test. This research was guided by the following two research questions: Do undergraduate science students using expert-created concept maps differ in their ability to enhance their recall of information about animal physiology when compared to students using visual organizers with limited hierarchy? How does prior knowledge affect the recall of students using concept maps and other visual organizers with limited hierarchy?

The data collected from the two groups was analyzed using regression analyses, ANOVA, and repeated-measures ANOVA. It was found that the hierarchical concept-mapping group grew more in their recall of information about animal physiology than the visual-organizer group [$F(1,38) = 7.70, p = .009$]. The results of these analyses were interpreted using the conceptual framework of cognitive load theory. This theory deals with the encumbrance on working memory that subsequently affects how one recalls information. The findings support the contention that hierarchical concept maps confer an advantage in the recall of science concepts when compared to visual organizers with limited hierarchy. This study lays the ground work for a doctoral study with 200 participants separated into four experimental groups ($n = 50$), with participants separated by high and low prior knowledge and the aforementioned visual organizers.
Acknowledgements

I would like to express my deepest gratitude to Richard Reeve, my supervisor, for providing me with all the necessary help I needed to complete this thesis. I would also like to thank William Egnatoff for agreeing to be a committee member and providing both wit and wisdom throughout my masters. Both professors gave considerable time and effort in every aspect of my thesis and I am very grateful to them. I am also grateful to Jamie Pyper for agreeing to be an examiner in my thesis defense.

I am indebted to John Kirby, a brilliant statistics professor, who selflessly supported the data analysis of this thesis. During my second semester, I dropped all of the classes that I was enrolled in at the time so I could take as many of his classes as possible. I wish to express my sincerest gratitude to William Higginson for providing me with constant edification and encouragement. Professor Higginson exhorted me to follow my wildest dreams and, oddly enough, a few of them came true. I also want to thank him for being one of the best teachers I have ever had. I would like to show my appreciation for Joan McDuff who is an excellent instructor with a contagious passion for teaching. I am grateful to the faculty, staff, and librarians of the Faculty of Education at Queen’s University. I would also like to show my appreciation for the Ontario Graduate Scholarship and Queen’s University for their financial support.

Finally, I would like to thank my friends and family. It is difficult to overstate my gratitude to Meena and Sou Thain for their unremitting love, understanding, and support. I am grateful to Brad, Chris, and Theressa for their magnanimous friendship. I would like to thank Paul and Saturlino for their unwavering faith in my abilities. I would like to express my gratitude for Tony and Joce for their company, kindness, and friendship.
Lastly, I would like to thank Nithum Thain for being an exceptional older brother and a constant pillar of support. To my friends and family, I love you all very much.
# Table of Contents

Abstract................................................................................................................................. ii  
Acknowledgements.................................................................................................................. iii  
List of Figures ........................................................................................................................... viii  
List of Tables ........................................................................................................................... ix  
Chapter 1 Introduction .............................................................................................................. 1  
  Overview ................................................................................................................................. 1  
  Definition of terms .................................................................................................................. 2  
  Research questions ................................................................................................................ 7  
  Potential significance of this study ......................................................................................... 8  
Chapter 2 Literature Review ................................................................................................... 10  
  Educational uses of concept maps ......................................................................................... 10  
  Research tool ........................................................................................................................ 10  
  Assessment tool .................................................................................................................... 12  
  Instructional tool ................................................................................................................... 13  
  Concept map construction ..................................................................................................... 16  
  Mixed-method research ......................................................................................................... 17  
  Collaborative maps .............................................................................................................. 18  
  Summary of literature review ............................................................................................... 20  
Chapter 3 Conceptual Framework: Cognitive Load Theory .................................................. 22  
  Introduction to cognitive load theory ..................................................................................... 22  
  Intrinsic load ........................................................................................................................ 22  
  Extraneous load .................................................................................................................... 23  
  Redundancy effect ............................................................................................................... 23  
  Split-attention effect ............................................................................................................. 25  
  Germane load ....................................................................................................................... 26  
  Cognitive load theory and the zone of proximal development ............................................. 27  
  Issues surrounding the use of concept maps ...................................................................... 28  
  Adaptability of concept maps .............................................................................................. 30  
  Summary of conceptual framework ..................................................................................... 32
Chapter 4 Methodology .................................................................................................................. 35
  Overview of the methodology ....................................................................................................... 35
  Participant selection ...................................................................................................................... 37
  Creation of multiple-choice test ................................................................................................... 39
  Multiple-choice test: Animal physiology ...................................................................................... 41
  Creation of maps .......................................................................................................................... 44
  Visual organizers ........................................................................................................................... 46
  Data collection ............................................................................................................................... 51
  Data analysis ................................................................................................................................ 52
  Alternative analysis ....................................................................................................................... 54
  Definition of variables .................................................................................................................. 54
  Ethics and incentives ...................................................................................................................... 56
  Summary of methodology ............................................................................................................ 57
Chapter 5 Results .............................................................................................................................. 59
  Overview of the results .................................................................................................................. 59
  Period I: Distribution of pre-test scores and post-test scores ....................................................... 60
  Period II: Distribution of pre-test scores and post-test scores .................................................... 61
  Descriptive statistics .................................................................................................................... 61
  Correlations between variables ................................................................................................... 64
  Distributions for periods I and II .................................................................................................... 67
  Regression analyses of individual outcomes in conjunction with map-type predicting post-test scores .......................................................................................................................................................... 68
  Repeated measures ANOVA ......................................................................................................... 71
  Repeated measures ANCOVA accounting for prior knowledge .................................................. 72
Chapter 6 Discussion ........................................................................................................................ 74
  Overview of the Discussion ............................................................................................................ 74
  Period I: Distribution of pre-test scores and post-test scores ....................................................... 75
  Correlations between predictors ................................................................................................... 76
  The benefit of hierarchy ................................................................................................................ 78
  Effects of prior knowledge ............................................................................................................ 81
  Controlling for the significant difference in pre-test scores ........................................................ 82
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive load in concept maps</td>
<td>85</td>
</tr>
<tr>
<td>Chapter 7 Conclusion</td>
<td>88</td>
</tr>
<tr>
<td>Limitations</td>
<td>89</td>
</tr>
<tr>
<td>Future impact of this study</td>
<td>91</td>
</tr>
<tr>
<td>References</td>
<td>94</td>
</tr>
<tr>
<td>Appendix A</td>
<td>102</td>
</tr>
<tr>
<td>Appendix B</td>
<td>103</td>
</tr>
<tr>
<td>Appendix C</td>
<td>105</td>
</tr>
<tr>
<td>Appendix D</td>
<td>107</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1: Animal physiology concept map. This concept map explains the concept of animal physiology with the intention of being used to answer the 18-question multiple-choice pre-test and post-test.................................................................49

Figure 2: Animal physiology visual organizer. This visual organizer with limited hierarchy explains the concept of animal physiology with the intention of being used to answer the 18-question multiple-choice pre-test and post-test...........................................50

Figure 3: Period I distribution scatter plot. This is a scatter plot depicting the pre-test and post-test score of the first 20 participants recruited in this study with a line of best fit ($R^2 = .029$, $p_s < .05$). The participants using the visual organizer with limited hierarchy are represented by blue dots and the participants using the concept map are represented by green triangles.................................................................62

Figure 4: Period I quartile distribution scatter plot. This is a scatter plot depicting the pre-test and post-test score of the first 20 participants recruited in this study. These participants are separated into quadrants based on the average score of the pre-test (7.85) and post-test (13.55).................................................................................................63

Figure 5: Period II distribution scatter plot. This is a scatter plot depicting the pre-test and post-test score of the second group of 20 participants recruited in this study with a line of best fit ($R^2 = .149$, $p_s < .05$). The participants using the visual organizer with limited hierarchy are represented by blue dots and the participants using the concept map are represented by green triangles.......................................................................................65

Figure 6: Period I and II distribution scatter plot. This is a scatter plot depicting the pre-test and post-test score of all 40 participants recruited in this study. The participants using the visual organizer with limited hierarchy are represented by blue dots and the participants using the concept map are represented by green triangles...............................................69
List of Tables

Table 1: Experimental Design ................................................................. 36
Table 2: Guiding Research Questions .......................................................... 37
Table 3: Descriptive statistics of participants mean pre-test score, post-test score, and academic average .......................................................... 64
Table 4: Correlation table of analyses variables ............................................. 66
Table 5: Independent regression analysis of each factor with map-type predicting post-test scores ........................................................................ 70
Table 6: Variance accounted for by outcomes in conjunction with map type predicting post-test score ........................................................................ 71
Chapter 1

Introduction

Overview

As educators, we must be able to develop the full potential of our students’ abilities if they are to realize the next generation of human achievement. This goal is impeded by difficulties in students’ recall of increasingly complicated academic subjects. It has been shown that students have significant difficulty in recalling scientific concepts because of such complexities (Johnstone, 1991). These difficulties stem partly from the way science is taught without consideration for students’ prior knowledge (Johnstone, 1991). Concept maps are a type of visual organizer that was created by Joseph Novak to advance the learning of science concepts by incorporating student prior knowledge (Novak, 1990).

The purpose of this thesis is to determine the efficacy of semi-hierarchical organization within concept maps in enhancing recall. In this study recall is defined as the ability to successfully answer questions about material that the student has previously encountered in an instructional intervention supported by different types of visual organizers. The results of the interventions will be interpreted through the conceptual framework of cognitive load theory. In this experiment, undergraduate students (N = 40) were recruited and randomly assigned to one of two experimental groups. These groups were asked to complete an 18-question multiple-choice test that measured their recall of
information about animal physiology, both before and after a visual organizer based intervention. The key terms are defined in the following section which will clarify the visual organizers used in this experiment, notable ideas in the literature review, and the conceptual framework.

Definition of terms

The following paragraphs explain key concepts as they are used in the thesis, the literature, and theory within which the thesis is framed.

Concept Maps. Concept maps are defined by Guastello, Beasley, and Sinatra (2000), as visual organizers that “make use of figures, lines, arrows, and spatial configurations to show how content ideas and concepts are organized and related” (p. 357). Propositions are statements that connect at least two nodes within a concept map through the use of linking words (Novak, 1990). Concept maps use propositions to identify the relationship between initially separated ideas. Figure 8 (see Appendix B) highlights the key structural components of concept maps described above.

Nodes. Nodes are the encircled ideas that represent the basic elements of a concept (Cañas et al., 2003). Nodes can be text, pictures, video, audio or any other relevant form of representing information that captures the anchoring ideas of a concept (Novak, 1993). Concept maps can be categorized as simple if they have a straightforward structure that contains a node-to-link ratio of approximately one. On the other hand, concept maps that are categorized as complex are circuitous in their structure and contain a node to link ratio that is notably larger than one (Meagher, 2009).
**Hierarchy.** In a literature review by Novak (2003) and other concept mapping experts (Cañas et al., 2003), it was determined that concept maps distinguish themselves from other visual organizers by virtue of their semi-hierarchical organization. Hierarchy is the arranging of ideas in a ranked order. Hierarchy allows ideas in a concept map to cluster around a central node, or a shared theme. Both hierarchy and propositions are considered by Novak (1993) to be integral features of concept maps. Since all ideas in a concept are connected in various ways, it is superfluous to identify all the possible connections in a concept map. Rather, what is important is to explain how major ideas in a concept relate to one another, hence the need for propositions to imbue these links with meaning.

Novak propagated Ausubel's concept of subsumption (Ausubel, 1968) in that “superordinate concepts subsume more specific, detailed concepts” (Cañas et al., 2003, p.13). When viewing a concept map it is assumed that general details would be closer to the topic node and context specific details located farther away from the topic node (Cañas et al., 2003). This trend can certainly be seen in most concept maps; however, other factors will determine the location of a node including visual clarity and cross-links (links between various lines of thought). As a result, it is said that concept maps contain a semi-hierarchical organization since their use and design does not strictly follow an absolute hierarchy (Cañas et al., 2003). Diagrams lacking these qualities are referred to under the broader category of visual organizers. This decision is made for the sake of clarity in nomenclature. It is yet to be determined if hierarchy is a necessary trait for
concept maps to be effective instructional tools (Mayer & Moreno, 2003; Rewey et al., 1989; Guastello et al., 2000; Katayama & Robinson, 2000).

**Working memory.** Working memory serves as the gateway between short-term memory, the memory that allows one to recall information presented within the past few seconds, and long-term memory, one’s permanent memory store. Working memory is limited to approximately four information elements, these elements are thought of as cognitive schemas (Sweller & Chandler, 1994). A schema is defined by Sweller (1994) as “a cognitive construct that organizes the elements of information according to the manner with which they will be dealt” (p. 296).

**Recall.** The process of retrieving information from long-term memory is referred to as recall. Tulving (2002) when discussing recall makes a distinction between two types, episodic and semantic. Episodic recall is the recall of personal information and past experiences (Tulving, 2002). Semantic recall is the recall of facts; this type of recall is more pertinent to this thesis.

In this thesis, an 18-question multiple-choice pre-test was used to determine the participants’ prior knowledge of animal physiology. This test determined what information on the subject of animal physiology was available to the students. Information that is considered available to the participants must be memorized in long-term memory, organized, and retrievable on a test (Tulving & Craik, 2000). The students were then provided with either a concept map or a visual organizer lacking hierarchy which contained all of the propositions associated with the concept of animal physiology.
It was assumed that this information had already been learned, but not all of it could be immediately retrieved from long-term memory (Tulving & Craik, 2000) without additional aid.

After a memory flushing exercise assumed to remove information from short-term memory received from the visual organizers, the students wrote a post-test (identical to the pre-test) that measured their recall of information about animal physiology. It was assumed that any improvements in test scores were due to improved access to information in long-term memory caused by studying the visual organizers (Tulving & Craik, 2000).

**Element-interactivity.** The degree of interaction between elements in working memory, referred to as element-interactivity, determines the difficulty of the learning task. Tasks that require a great deal of interaction between elements are said to have high element-interactivity and are considered of greater difficulty than learning tasks that require low element-interactivity (Schnotz & Kürschner, 2007). The following example, taken from an article by Sweller (1994), clearly outlines the difference between high and low element-interactivity.

Low-element interactivity tasks allow elements to be learned serially rather than simultaneously. The tasks can be fully understood and learned without holding more than a few elements in working memory at a time. High element-interactivity tasks are at the other end of the continuum of complexity. To understand and learn these tasks, several elements must be manipulated in working memory simultaneously. For example, learning
the vocabulary of a foreign language is a low element-interactivity task, whereas, learning the grammatical properties is a high element-interactivity task because the elements interact and learning them as individual elements may be meaningless. As an example, we only can understand the required order of words in the English language by considering all of the words in a phrase. It would make little sense to consider one word at a time. Knowing that all of the words in a phrase are in the correct order cannot be learned by considering each word individually. Failing to relate each word to the others results in failure to learn the task. Learning word order is a high element-interactivity task because all of the elements must be processed in working memory simultaneously.

The virtue of concept mapping is that it allows the learner to store ideas and connections in the instructional tool and this alleviates the burden of storing these components in working memory (Edwards & Fraser, 1983). Learners are no longer tasked with having to store all the possible connections between ideas contiguously. Instead, when using a concept map, the learner need only make one connection at a time; this extends the concept map by one node, link or linking term, and by replicating this straightforward process a complex concept map can be created (Anderson-Inman & Horney, 1997).

Cognitive load theory. The conceptual framework for this thesis is cognitive load theory (See Chapter 3), which explores the encumbrances on our working memory that subsequently affects the efficiency with which we learn information (Sweller, 1994). Cognitive load has been defined as “a multidimensional construct representing the load
that performing a particular task imposes on the learner’s cognitive system” (Paas et al., 2003, p. 64). This theory suggests that we have a limited working memory that serves as the gateway to our long-term memory (Kirschner, 2002). This theory outlines three categories of cognitive load; intrinsic, extraneous, and germane (Sweller, 1994; Schnorr & Kürschner, 2007). Intrinsic load is the inherent complexity of a cognitive task and this difficulty is thought to be unalterable (Kirschner, 2002). Extraneous load is an unnecessary cognitive load that is caused by improper instructional design. This inefficiency is brought about when information that does not contribute to the understanding of a concept is added to the instruction of that concept (Schnorr & Kürschner, 2007). Lastly, germane load is cognitive load that does not directly reflect the learning outcomes surrounding a concept and yet aids in the learning of that information (Sweller et al., 1998).

**Research questions**

In this thesis, the efficacy of concept maps in enhancing recall by increasing germane cognitive load and decreasing extraneous cognitive load was explored. The following questions guided this research:

1. Do undergraduate science students using expert-created concept maps differ in their ability to enhance their recall of information about animal physiology when compared to students using visual organizers with limited hierarchy?
2. How does prior knowledge affect the recall of students using concept maps and other visual organizers with limited hierarchy?
Potential significance of this study

This research has value in both academic and classroom settings. This research is useful in academia because it contributes to the current literature on both cognitive load and concept mapping theory. Currently, little research addresses how the constraints imposed by concept maps alter a specific type of cognitive load. This study is novel in so far as it systematically controls one characteristic of concept maps to quantifiably determine its value in the recall of information.

Visual organizers are considered to be underutilized in classrooms despite their usefulness (Guastello et al., 2000). Understanding the features of concept maps and their contribution to recall, would be beneficial in creating more efficient visual organizers. Furthermore, this research re-examines the value of concept maps by showing that the additional constraints imposed by this instruction may only serve to increase extraneous load and therefore may actually reduce recollection of the concept.

Although current research is typically supportive of concept map interventions (Meagher, 2009; Leutner et al., 2009; Guastello et al., 2000; Cooper et al., 2001), not much is known about the precise way in which these interventions are beneficial. After completing a meta-analysis of 55 concept map interventions, Nesbit and Adesope (2006) declare the future of this research needs “to explain how students learn from concept maps… research is needed that compares the use of concept maps, graphic organizers and outlines” (p. 435). It would be valuable for future research to focus on controlled experiments that would methodically test various constraints of concept maps, to
understand if these constraints are meaningful or an unnecessary encumbrance (Katayama & Robinson, 2000).
Chapter 2

Literature Review

Educational uses of concept maps

The literature surrounding the educational uses of concept maps (Novak, 1990; Wheeldon, 2010; Meagher, 2009; Royer & Royer, 2004; Guastello et al., 2000) falls into three broad categories according to their use. Concept maps can be used as a research tool, an assessment tool, or as an instructional tool (Nesbit & Adesope, 2006). This chapter examines and summarizes the current literature surrounding concept maps, initially explaining the three main uses. Then this chapter reviews the literature explaining the construction of these maps, the value of concept maps in mixed-method research, and collaborative mapping. This literature informs the methodology of this study and helps to clarify its purpose. Although this thesis is primarily concerned with the use of concept maps as an instructional tool, a concept map was used as a research tool that aided in the writing of this literature review (see Appendix A).

Research tool

Concept maps were originally conceived by Joseph Novak as a research tool that would allow researchers to gain insight into students’ conception of scientific knowledge (Novak, 1990). Edwards and Fraser (1983) claim that using concept maps to this end has been found to be as effective as conducting an interview. Researchers using concept maps and other visual organizers have in effect been able to peer into the cognitive processes of
students (Novak, 1993; Mason, 1992). In one such study, the authors, Hmelo-Silver and Pfeffer (2004), examined the differences between experts’ and novices’ understanding of complex systems. This mixed-method study observed differences in understanding of aquaria among grade seven students, eleven pre-service teachers, and eight experts. The eight experts were further subdivided into two groups, hobbyists who had spent between 10 and 30 years maintaining numerous aquaria and biologists who had an advanced degree in the subject as well as research publications (Hmelo-Silver & Pfeffer, 2004). The participants were asked open-ended questions about aquaria. The participants were then asked to respond to these questions with both a diagrammatic and oral responses. These answers were then coded in a structure, behavior, and function (SBF) framework, which determined whether the answers given fit into one of these three categories.

Quantitative analysis yielded a significant difference between the expert group and the two novice groups in terms of interactions between the structure, behavior, and function of the concepts. This analysis also showed a significant increase in experts’ use of behavioral and functional concepts; however, there was no significant difference in the use of structural concepts. Qualitative analysis was conducted on the transcript of the interviews which supported the quantitative analysis (Hmelo-Silver & Pfeffer, 2004). This indicated further differences between experts and novices in their understanding of this complex system.

This study highlights the effectiveness of concept maps as a research tool that can advance our current understanding of the recollection process. Experts had explanations
for their concepts that were more elaborate and better integrated than the explanations provided by the novices. Novices were more likely to be fixated on the physical aspect of a system; since this part of a system is tangible, it is easier to understand (Hmelo-Silver & Pfeffer, 2004). Furthermore, differences were found between the two groups of experts: members of the biologist group used explanations that contained abstract concepts; whereas, members of the hobbyist group used explanations that were concrete and situated to pragmatic examples (Hmelo-Silver & Pfeffer, 2004).

Assessment tool

As stated by Novak and Gowin (1984), “Concept mapping’s most significant contribution to the improvement of educating may be the ultimate improvement of evaluation techniques” (p. 23). Concept maps have been used primarily for formative assessment, an interim assessment that facilitates student learning. Students are asked to create concept maps to demonstrate their understanding of a topic or their solution to a problem. These concept maps are usually scored in one of two ways, the first method envisioned by Novak is correlating the understanding of a topic with the complexity of the map that is derived (Wheeldon, 2010). This is done by using a scoring system that quantifies the structural components of a concept map (valid propositions, level of hierarchy, degree of branching and cross-links) (Wheeldon, 2010). Complexity found in the assessment of student constructed concept maps is considered to be indicative of learning (Meagher, 2009). This assumption seems credible since studies have been able to correlate greater complexity in concept maps with greater achievement in assessments.
(Royer & Royer, 2004; Guastello et al., 2000). This method only uses an expert in the capacity of determining whether or not a proposition is valid.

The second method of scoring relies far more on experts’ conceptions of the topic (Ruiz-Primo & Shavelson, 1996). This method compares the students’ concept map directly to that of an expert, which makes us consider the expert’s concept map as some ideal that we wish to emulate. These two markedly different forms of assessment demonstrate the adaptability with which concept mapping techniques can be utilized, a point that is discussed further when addressing the conceptual framework of this thesis.

**Instructional tool**

The most relevant use of concept maps with regards to this thesis is the use of concept maps as an instructional tool. Teachers may use concept maps as a visual organizer that clearly demonstrates the relationships among concepts. Ausubel believed advanced organizers could facilitate learning by providing background knowledge that could provide a meaningful context for new knowledge (Novak, 2010). Various studies have reported the benefits of employing concept mapping instruction (Guastello et al., 2000; Shen et al., 2007; Horton et al., 1993; Mason, 1992). These studies did show improvement in student learning when the visual organizer was employed; however, they did not clearly define the traits of a concept map and therefore ambiguity exists concerning what specific features makes this tool effective.

In one such study that aims to improve science content recollection for students with low prior knowledge, Guastello, Beasley, and Sinatra (2000) engaged in a
quantitative study to test the effectiveness of an intervention using concept-mapping instruction. The participants of this study were 142 grade seven students in an urban parochial school located in Brooklyn, New York. These students were randomly assigned into two groups; one group was taught a lesson using the “read-and-discuss, teacher directed method” (Guastello et al. 2000, p. 356), while the other group was given an identical type of introductory lesson and an additional concept mapping intervention. The participants wrote two tests prior to any instruction; the first was the Comprehensive Assessment Program (CAP), which ensured the children were in fact academically low-achieving. Guastello et al. (2000) claim “the internal consistency reliability of the various components of the CAP is reported to range between the high .80s and mid-.90s” (p. 359). This test revealed that 23 students scored above grade level on the CAP test and were subsequently disqualified from the study. The remaining 127 participants were divided into two groups, either an experimental or control group, and no significant difference was found between these groups with regards to their CAP scores. The second assessment was a criterion referenced-test, which consisted of 20 questions, based on the specific chapter the students were studying. It is worth noting that the control group in this study was not provided any visual organizer as a substitute for concept maps.

An analysis of covariance demonstrated “a strong and statistically significant treatment main effect, (F( 1, 121) = 1,261.56, p <.0001), favoring the experimental group” (Guastello et al., 2000, p. 362). A pooled variance estimate was conducted and the effect size for the treatment was estimated to be approximately 5.8 standard deviations.
The authors concluded from this study that the concept mapping treatment was responsible, in this instance, for the almost six standard deviation increase in learning when compared to a traditional lecture (Guastello et al., 2000).

A major design flaw in the aforementioned study by Guastello, is that the concept maps used for instructing students did not make use of propositions. Propositions are essential for an effective concept map; an increase in proposition use, according to Meagher (2009), is considered to be an increase in learning. Even though propositions were missing from this intervention it was still able to produce a six standard deviation difference between the experimental and control group.

This study supports the notion that visual organizers can be effective instructional tools for traditionally low achieving elementary students. However, the authors erroneously conclude that concept maps confer this benefit. The visual organizer used in the study cannot aptly be characterized as a concept map and this study may accidentally exaggerate the benefit of this tool. It is possible that the success of this intervention hinges on the simplicity of the visual organizer employed. Perhaps the constraints of using an authentic concept map, possibly an additional extraneous load, would only obfuscate complex scientific concepts.

The exceptionally large difference in standard deviation is likely caused by the control group attempting to learn a rather complicated scientific concept without any support from images or diagrams. This issue may be exacerbated if the final assessment demands or encourages students to answer questions with a diagram. Instead of the
purported advantage in learning, students in the experimental group may just be provided with ideal answers for the test, an advantage not shared by the control group. Unfortunately, the authors of the study did not provide sufficient details about the assessment to properly resolve this issue (Guastello et al., 2000).

**Concept map construction**

Concept maps are usually created in one of two ways either using paper and conventional writing utensils (pencils) or through computer software. In an experimental study comparing the differences between the two methods of creating concept maps, Royer and Royer (2004) discovered that computer-generated maps created more complex maps than paper/pencil maps. This study involved 52 participants, who were students in a combined Grade 9/10 biology class. These students were divided into two treatment groups: one group of students was asked to create maps using paper/pencil, the other group was instructed to use *Inspiration©* software.

The mean difference in complexity between the two groups was calculated using t-test analysis. It was found that the computer-generated maps were significantly more complex than the paper/pencil generated maps. As previously mentioned in the research uses section, an increase in complexity of concept maps is considered to be indicative of an increase in learning (Meagher, 2009). Computers also conferred other benefits when creating concept maps, such as ease of manipulation, dynamic linking, and the ability to make revisions (Royer & Royer, 2004). This is valuable since concept maps were originally intended to be organic documents that are constantly reconfigured.
Students cited a preference, in both interviews and a final survey, for using computers when generating their concept maps (Royer & Royer, 2004). Their preferences were attributed to greater ease of use and clarity in presentation (Royer & Royer, 2004). The two visual organizers required for this study were created using computers since it ensures that the maps are legible and consistent for all the participants in the study. Using software maintained consistency in the visual clarity of the nodes between the maps. This consistency increased the trustworthiness of this study and enhanced the validity of these findings.

**Mixed-method research**

Concept maps are considered especially valuable when conducting mixed-methods research and can provide “an accessible and innovative means to combine the clarity of quantitative counts with the nuance and perspective of qualitative reflection” (Wheeldon, 2010, p. 88). Since the individual aspects of the concept maps are clearly labeled, there is a consistency between maps that lends an objectivity that is required for effective quantitative analysis (Morgan, 2007). Anchoring ideas within concept maps are encircled in nodes, and are connected by linking words to create propositions. Therefore, researchers need not infer between the relationships of nodes in a participant’s response. An abundance of concept mapping software makes objective analyses straightforward since there is pronounced consistency between the components, and quantitative tools are often included in these programs (Royer & Royer, 2004). Concept maps consist of a
series of propositions that provide a depth of language and personal insight that are conducive to rigorous qualitative analysis.

**Collaborative maps**

Collaborative concept mapping is defined as “a process where two or more students are engaged in coordinated and sustained efforts in the creation of one or more concept maps to learn and construct knowledge” (Shen et al., 2007, p. 480). Collaboration has become increasingly important in the modern educational ethos. There is evidence to suggest that collaboration, under the right circumstances, can enhance the effectiveness of interventions like concept mapping (Huff & Jenkins, 2002; Shen et al., 2007; Fraser & Novak, 1998). In a literature review by Shen, Turner and Gao (2007), it was found that collaborative concept mapping interventions have contradictory findings. Their review used studies published from the past 20 years in which concept mapping was used either as an instructional or assessment tool. The review summarized major studies and made it explicitly clear that it would present both the negative and positive findings of the interventions. It concludes by suggesting how collaborative and individual concept mapping instruction can be improved by supporting one another.

The review by Shen *et al.* (2007), presented many studies that supported the notion that collaborative concept mapping produced better results than individual concept mapping exercises. Collaborative concept mapping tended to show more intricate details, more linkages and better improvements in declarative knowledge than individual concept mapping (Shen et al., 2007). Furthermore, concept maps made collaboratively could be
used as a “referential anchor”, an artifact that would support mutual understanding between the collaborating parties (p. 482).

However, there are contradictory studies that present evidence that collaborative concept mapping is less effective than individual concept mapping (Chiu, 2003). Many interventions with individualized concept mapping instruction have been successful (Horton et al., 1993; Mason, 1992). A great deal of time in collaborative concept mapping interventions went towards process-oriented exchanges. These exchanges were meant to organize the collaborative effort and to delegate responsibility (Chiu, 2003). Although these exchanges were often productive and may have been beneficial for learning about collaboration, they are counter to the original *raison d’être* of the intervention (Novak, 2002). Ultimately, these processes detracted from the total time spent towards learning the concept.

In a literature review by Novak and other concept mapping experts (Cañas et al., 2003), it was suggested that:

In facilitated mapping sessions, the skill of the facilitator has a significant bearing on the quality of the result. If a session involves brainstorming, it is imperative that the facilitator has good keyboarding skills in addition to an understanding of how to facilitate the building of concept maps. It is also quite clear that group dynamics play a significant role in the quality of a concept mapping session. It can be less than optimal, or even counter-productive for large, highly contentious groups of people to attempt to create collaborative concept maps (p. 64).
Other studies found that collaborative concept mapping did not have any effect on students’ self-efficacy (Shen et al., 2007; Ledger, 2003). Collaborative concept mapping also lead to incorrect notions being unchallenged by the group; since often times, the entire group depended on a few individuals whose assertions went unchallenged (Roth & Roychoudhury, 1992). These negative finding do not preclude the use of collaborative concept mapping but instead require collaborative concept mapping to be cautiously implemented (Shen et al., 2007). It would seem that collaborative concept mapping is optimized when preceded by individual work (Chiu, 2003). This type of preparation allows students to begin thinking as individuals and reduces the tendency for the entire group to depend on the thinking of a few.

**Summary of literature review**

The current literature surrounding cognitive load theory speaks often and favorably on the role of visual organizers, and their ability to enhance the learning of science concepts (Schnotz & Kürschner, 2007; Mayer & Moreno, 2003; Ruiz-Primo & Shavelson, 1996; Shen et al., 2007; Royer & Royer, 2004; McClure et al., 1999). However, the issue of concept maps being ill-defined in the reporting of interventions using visual organizers still persists. Often studies (Mayer & Moreno, 2003; Rewey et al., 1989) use concept maps while using the umbrella term of visual organizers to describe them. Other studies (Guastello et al., 2000; Katayama & Robinson, 2000) state that their instruction employs a concept map when in fact it uses a different type of visual
organizer. These studies will be useful in informing the methodology of this thesis in Chapter 4.
Chapter 3

Conceptual Framework: Cognitive Load Theory

Introduction to cognitive load theory

Nesbit and Adesope (2006) conducted a meta-analysis in which they chronicled the history of concept mapping literature. The authors mention a few concept mapping studies that speak directly to the cognitive load of the student (Hall, Hall, & Saling, 1999; Rewey et al., 1989). Cognitive load theory serves as the conceptual framework for this study as it may explain the effects of semi-hierarchical organization in the process of recall when using visual organizers. As previously mentioned (see Definition of terms), cognitive load has been defined as “a multidimensional construct representing the load that performing a particular task imposes on the learner’s cognitive system” (Paas et al., 2003, p. 64). Cognitive load theory explores the encumbrances on our working memory that subsequently affects the efficiency with which we learn and recall information. This theory outlines three categories of cognitive load: intrinsic, extraneous, and germane (Sweller, 1994; Schnitz & Kürschner, 2007).

Intrinsic load

Intrinsic load is the inherent complexity of a cognitive task and this difficulty is thought to be unalterable (Kirschner, 2002). However, what comprises intrinsic load will vary depending on the differences in learning outcomes between students. For example, a student in law school may be required to know the names of legal cases and precedents
when discussing political issues in the classroom. A student in another faculty, for instance sociology, would not be required to know these precedents even if they were to discuss the same social issue as the aforementioned law student. Therefore, information that is considered intrinsic load for a student in law school may be extraneous or germane load for a student in another faculty; by virtue of the fact that the learning outcomes are different between these faculties. The archaic notion of instruction being subsumed by a single type of cognitive load is considered outdated (Schnotz & Kürschner, 2007). Instead, a nuanced approach to cognitive load theory has recently arisen which determines the cognitive load at play by virtue of the agent and the circumstances surrounding the learning task (Schnotz & Kürschner, 2007).

**Extraneous load**

Extraneous load is an unnecessary cognitive load that is caused by improper instructional design. This inefficiency is brought about when information that does not contribute to the learning of a concept is added to the instruction of that concept (Schnotz & Kürschner, 2007). Two specific types of extraneous load that are discussed in this thesis are redundancy effect (Sweller & Chandler, 1994) and split-attention effect (Mayer & Moreno, 2003). Instructors should aim to reduce as much extraneous load in their instruction as they can.

**Redundancy effect**

Redundancy effect refers to the beneficial effect of removing redundant information. When a student is presented with information that is already learned, that
student performs worse at a learning task than a similar student who does not have to process redundant information (Sweller & Chandler, 1994). Redundancy effect is explained in cognitive load theory by the limitations imposed on learning by working memory.

The degree of interaction between elements in working memory, referred to as element-interactivity, determines the difficulty of the learning task. As previously mentioned, the limitation of four interacting information elements constrains the amount of information that can be readily memorized and stored in long-term memory (Sweller & Chandler, 1994). If a student is forced to process and comprehend information that is already stored in long-term memory, their working memory is unnecessarily occupied with elements that are not aiding the acquisition of new knowledge (Schnotz & Kürschner, 2007). This causes the student to try to once again memorize information that is already stored in long-term memory, as well as reducing the available working memory.

A concept map is intended to be a skeletal representation of a concept without additional information. A well-designed map allows its reader to quickly and efficiently understand the visual information that is being presented. This is beneficial not only in reducing element interactivity but also in limiting the redundancy present in the map. Since nodes represent anchoring ideas it is unlikely that a concept map will present the same idea repeatedly (Anderson-Inman & Horney, 1997). Instead, a concept map will make multiple links to one node allowing the student to view complex interactions.
between ideas, without using redundant information that would be necessary if the instructional format depended entirely on text (McClure et al., 1999). By reducing the redundancy in instruction, the student can acquire more information with the same effort.

**Split-attention effect**

Another form of extraneous load, split-attention effect, is caused when a student must integrate multiple sources of visual information because it is necessary for the understanding of the concept as a whole (Sweller, 1994). Since these visual components could be integrated in the instruction, the mind is wastefully committing resources in an attempt to make sense of this information. This inefficiency can be rectified by integrating the visual sources, combining them so they can be seen in one view, thus relieving the burden of integration from working memory (Mayer & Moreno, 2003).

Concept maps allow a student to view a concept as a whole without using additional working memory to bring forth these ideas (Anderson-Inman & Horney, 1997). The use of text in concept maps is concise, thus allowing students to quickly search for a specific idea. Since linking words are required to be present between nodes; text becomes naturally aligned with images and other structures within concept maps (Mayer & Moreno, 2003). The relationships between these ideas (propositions) are brought forth as a result of this alignment. Ultimately, concept maps have an implicit propensity for alignment; therefore working memory is no longer needed to integrate the images and this resolves the split-attention effect (Mayer & Moreno, 2003).
**Germane load**

Prior to the mid-1990s, cognitive load theorists were predominantly interested in reducing extraneous cognitive load in instructional materials (Mayer & Moreno, 2003). At that time, scholars believed intrinsic load was unchangeable and extraneous load was the only load that was mutable. This focus did not change until Paas and van Merriënboer (1994), while studying the effects of cognitive load in problem solving, theorized about a new type of load (germane) associated with positive learning outcomes.

Germane load is cognitive load that does not directly reflect the learning outcomes of a learning task and yet aids in the learning of that material (Sweller et al., 1998). Learned material that is transferred from working memory into long-term memory is stored in higher order units called cognitive schemata (Schnotz & Kürschner, 2007). As learning progresses, elements that were originally separated become increasingly bound forming a unified concept. This compression of elements into schema is referred to as schema acquisition and this process reduces the cognitive load necessary for the utilization of these elements in the process of learning information (Sweller & Chandler 1994; Sweller et al., 1998). Eventually, a schema can be held in working memory without using any active effort in which case schema automation is said to have occurred (Sweller et al., 1998).

The purpose of this thesis is to determine the efficacy of semi-hierarchical organization within concept maps in enhancing the recall of information about animal physiology. Concept maps are effective, as would any visual organizer, at reducing some of the cognitive burden of recalling a science concept (Cooper et al., 2001). However,
specific characteristics of concept maps (alignment, simplification of concepts and adaptability) help expand this initial benefit; these characteristics will be discussed further in the remainder of this chapter. Concept maps have additional constraints, compared to most other visual organizers, which seem to limit their use as an instructional tool. These constraints add to the cognitive load of students; however, if these constraints are ultimately beneficial they constitute germane load. As we will see in the following sections one should tailor the cognitive load of the instruction to the students’ prior knowledge. Concept maps can be of greater benefit than visual organizers in recalling science concepts if the characteristics of the concept maps are utilized properly.

**Cognitive load theory and the zone of proximal development**

Recent work by Schnotz and Kürschner (2007) has focused on reconsidering cognitive load theory from the perspective of Vygotsky’s zone of proximal development. Vygotsky understood learning as a process that required a balance between comfort and challenge. Ideally, a learner should be taught in their zone of proximal development, the range between the lower and upper boundary of task difficulty. Schnotz and Kürschner (2007) defined these limits as follows, “the lower limit…is defined as the most difficult task the learner can perform successfully without help, whereas the upper limit… is defined as the most difficult task the learner can perform successfully with help” (p. 487).

In one respect, a learner who is not adequately challenged is often unmotivated to learn (Schnotz & Kürschner, 2007). Furthermore, learning is contingent on the
acquisition of novel information and learners who are below their zone of proximal
development are not addressing enough new information to learn in great depth. On the
other hand, learners who are well above their zone of proximal development are unable to
learn because the material is excessively challenging and making great strides in learning
becomes implausible (Schnitz & Küschner, 2007).

**Issues surrounding the use of concept maps**

Ruiz-Primo and Shavelson (1996) mention possible issues related to the use of
concept maps that may in fact increase the extraneous load rather than the germane load
for students. Concept maps can be idiosyncratic and if the student does not create the
concept map then the diagram may add unnecessary encumbrances when attempting to
understand the creator’s meaning. Initially, it would seem that this situation could be
rectified by having the student make their own concept map; however, this is not feasible
for many students with low prior-knowledge (Amadieu et al., 2009).

Additionally, students who have low prior-knowledge, who may struggle with
deeper forms of learning, would need time to adjust to using concept maps (Ruiz-Primo
& Shavelson, 1996). This is because concept maps are organized primarily to enhance
meaning as opposed to aesthetic appeal. Organization of knowledge based on meaning is
conducive to deeper approaches to learning rather than memorizing the information based
on shallow memory cues, like mnemonic devices (Ruiz-Primo & Shavelson, 1996).

The previously mentioned study by Guastello, Beasley, and Sinatra (2000) seems
to contradict the previous sentiment, and this study provides evidence that concept maps
are exceedingly beneficial for low prior knowledge students. This is perhaps not a comprehensive response to the criticism since the visual organizer in the study was missing linking words and propositions, which are integral parts of a concept map. Furthermore, if this visual organizer were to be considered a concept map it would have been characterized as a simple branched tree map with highly directed instruction (Ruiz-Primo & Shavelson, 1996). This incomplete map and highly directed instruction would require little cognitive load to properly utilize, and one could make the argument that a more complicated map would have been considerably less effective for low prior knowledge students.

Some of these criticisms have merit theoretically; but, an exhaustive search of the literature has yet to yield any instructional intervention that has found problems explicitly and adversely affecting student achievement (Guastello et al., 2000; Meagher, 2009; Leutner et al., 2009; Cooper et al., 2001; Cañas et al., 2003; Amadieu et al., 2009). After completing a meta-analysis of 55 concept mapping interventions, Nesbit and Adesope (2006) found:

No categories or conditions in which concept maps produced significant negative effects, and, aside from the theoretical objection one might pose that frequent use of concept maps could reduce practice in reading and writing text, no potentially detrimental effects have been identified (p. 435).
Concept maps are, however, scrutinized for their lack of clarity with regard to assessment practices. Ruiz-Primo and Shavelson (1996) insisted that an “integrative working cognitive theory is needed to begin to limit variation in concept mapping techniques for assessment purposes” (p. 569). Perhaps this problem can be rectified by utilizing current advances in cognitive load theory to construct a framework that would enhance assessment validity and reliability. However, this is a predicament for concept mapping interventions at large and does not apply to the narrower scope of this thesis which is the use of concept maps as an instructional tool.

**Adaptability of concept maps**

Concept mapping instruction can be employed with varying degrees of directedness making it beneficial for teaching a group of students with differences in prior knowledge. Students can be asked to build their own concept maps: this type of instruction is said to have a low degree of directedness (Ruiz-Primo et al., 2001). When maps are created by the students they will contain more personalized connections with the material, increasing the germane load from the lesson. However, this type of instructional format will increase the overall cognitive load and is therefore ideal for students with high prior knowledge (Ruiz-Primo et al., 2001).

Alternatively, concept maps can be highly directed requiring less cognitive load on the part of the learner (Ruiz-Primo et al., 2001). In this method of instruction, the instructor provides students with an expert map that, following a behaviorist framework encourages students to memorize the ideal form of the concept. A teacher may also opt
for a moderate degree of directedness by providing some aspects of a concept map while asking the students to modify the rest of the map as they see fit (Ruiz-Primo et al., 2001).

Concept mapping instruction has even greater adaptability when considering the various types of maps that can be created (Meagher, 2009). This claim was investigated in a study that explored the effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning from non-linear documents (Amadieu et al., 2009). In this study 24 adults, studying HIV at the university level, were randomly assigned into one of two groups that were divided by the complexity of the concept map they would utilize. One group was given instructional material that was presented on a hierarchical concept map. The other group was given a network map that is considerably more complex and required greater cognitive load to comprehend. These two groups were then subdivided into students with high prior knowledge and students with low prior knowledge. Since the number of participants involved in the study was quite small, nonparametric tests were used to analyze the data, with the significance level set at .05 (Amadieu et al., 2009). A Mann-Whitney U test was conducted and the difference in prior knowledge between the low and high prior knowledge groups was found to be significant, U(12, 12) = 2, \( p < .001 \). Students with low prior knowledge when tested for advances in conceptual knowledge benefited from using the simpler concept map rather than the more complex network concept map, the difference in mean achievement was found to be significant, U(6, 6) = 7, \( p = .036 \).
Students with low prior knowledge also reported exerting more mental effort in answering questions than did the high prior knowledge group irrespective of the map used (Amadieu et al., 2009). Furthermore, students with low prior knowledge were found to have greater disorientation when using the network map $U(6, 6) = 2, p = .005$. Students with high prior knowledge did not have a significant difference in disorientation when using the different concept maps $U(6, 6) = 8, p = .054$, however the test was very close to being statistically significant.

Concept mapping techniques can be employed with a range of complexity making them beneficial for teaching a group of students with disparities in prior knowledge. This study quantifies the value in matching concept map complexity to student prior knowledge (Amadieu et al., 2009). Needless to say, this study would have benefited substantially from a greater number of participants, given that it was primarily quantitative in nature. However, this study serves to buttress the notion that concept maps are considerably more effective when concept map complexity is matched appropriately with student prior knowledge.

**Summary of conceptual framework**

This thesis addresses the efficacy of concept maps in enhancing the recall of information about animal physiology using the framework of cognitive load theory. Current educational literature supports the notion that the instructional use of concept maps affects students’ cognitive load favorably (Amadieu et al., 2009; Ruiz-Primo et al., 2001; Schnottz & Kürschner, 2007; Mayer & Moreno, 2003). Concept maps are effective,
as would any visual organizer, at reducing some of the cognitive burden of recalling a
science concept (Cooper et al., 2001), but specific characteristics of concept maps
(alignment, simplification of concepts and adaptability) help expand this initial benefit.

The nodes within concept maps are well aligned with each other as a result of
their efficient design, and their use helps reduce split attention (Mayer & Moreno, 2003).
Concept maps simplify concepts in so far as they present a skeletal representation of
concepts without additional information. Anchoring ideas within a concept are
represented by nodes and these nodes are seldom repeated, thus redundancy is unlikely to
occur (Novak, 2002; McClure et al., 1999). Finally, concept maps can be employed with
an adaptable range of directedness and complexity (Ruiz-Primo et al., 2001). This allows
teachers to better match difficulty of instruction with students’ zone of proximal
development (Schnotz & Kürschner, 2007).

This is not to imply that concept maps are without flaws. Complex concept maps
are considered to be excessively taxing for students with low prior knowledge (Ruiz-
Primo & Shavelson, 1996). As previously mentioned, current research is typically
supportive of concept mapping interventions (Meagher, 2009; Leutner et al., 2009;
Guastello et al., 2000; Cooper et al., 2001); however not much is known about the precise
way in which these interventions are beneficial. It would be valuable for future research
to focus on controlled experiments that would methodically test various components of
concept maps to understand if these constraints are meaningful or an unnecessary
encumbrance (Katayama & Robinson, 2000; Nesbit & Adesope, 2006). The purpose of
this thesis is to determine the efficacy of semi-hierarchical organization within concept maps in enhancing the recall of information about animal physiology.
Chapter 4

Methodology

Overview of the methodology

This study focused on university students’ recall of information about animal physiology. In this study two experimental groups were created. One group was presented with a concept map that included hierarchy, and the other presented with a comparable visual organizer that has limited hierarchy. Each experimental group consisted of 20 university students enrolled in an undergraduate science program with the exception of four bachelors of education students (see Table 1). One-way ANOVA analysis determined possible differences in recollection of the visual organizers between the two groups. Further analysis was conducted on the change in score of the pre-test and post-test multiple-choice assessment. The following questions along with the accompanying statistical analysis presented in Table 2 guide this research.
<table>
<thead>
<tr>
<th>Group</th>
<th>Assignment</th>
<th>n</th>
<th>Pre-test</th>
<th>Treatment</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Random</td>
<td>20</td>
<td>18-QMCT*</td>
<td>Concept Map</td>
<td>18-QMCT</td>
</tr>
<tr>
<td>2</td>
<td>Random</td>
<td>20</td>
<td>18-QMCT</td>
<td>Visual organizer with limited hierarchy</td>
<td>18-QMCT</td>
</tr>
</tbody>
</table>

*Note: *18-question multiple-choice test

\*N = 40
Table 2: Guiding Research Questions

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Statistical Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do undergraduate science students using expert-created concept maps differ in their ability to enhance their recall of information about animal physiology when compared to students using visual organizers with limited hierarchy?</td>
<td>ANOVA Regression Analysis Repeated Measures ANOVA</td>
</tr>
<tr>
<td>How does prior knowledge affect the recall of students using concept maps and other visual organizers with limited hierarchy?</td>
<td>ANCOVA Regression Analysis Repeated Measures ANCOVA</td>
</tr>
</tbody>
</table>

**Participant selection**

Undergraduate and bachelors of education students enrolled in biology courses were asked to join this experiment in two periods via the university’s LISTSERV (email list), poster advertisements and classroom presentations. Four bachelors of education students who had recently completed their undergraduate degree in biology participated in this study in period I and they were considered biology students in the subsequent analyses. Approximately 1000 different students were approached to join this study in classroom presentations and a total of 3000 students were contacted via email. The estimated number of accepted participants was expected to be at least 30, and there was no limiting age, gender or prior knowledge for participants in this study. Unfortunately, after the first period of participant recruitment only 20 participants were found who successfully completed this study; all of these participants were unpaid. In this first period, 10 participants were assigned to each of the two groups. Eventually, this study did
engage in a second period of participant recruitment where students were paid $15 for their participation. During the second period of recruitment another 20 participants were gained contributing to a total participant pool of 40 students.

In both periods, once the students were part of the participant pool, they were randomly assigned to one of two groups: visual organizer with limited hierarchy group or the concept mapping group (with hierarchy). This was done by assigning each student a number and using a non-repeating random number generator to select students for each group. Students were then asked to join a single 45 minute experimental session and they were given several different dates for their convenience. During period I, these sessions were held during the evening between 5:30 – 7:30 pm to mitigate possible environmental biases, such as a difference in alertness based on the time of day. All the sessions were also held in the same location, a classroom with the same configuration for every session. This classroom was located in a building on the periphery of campus, which is reasonably far from their regular class location. Students selected for this study were given a letter of information and two consent forms, one to be completed and given to the researcher, and one that they would take with them.

During the first period of this study, the initial number of participants limited the statistical significance of the results, and ultimately diminished the potential validity of the study. The initial data collection period of one week was extended to a month but I managed to only recruit 20 students. This was due to a confluence of factors including the
time of year, the students I was trying to recruit, the location of the study and the inability
to provide a monetary reward for participation.

A month later, period II of the recruitment was initiated because of the low
number of participants. The recruitment method changed slightly offering an additional
incentive for participants, $15 (CDN), for their participation. Students who had
participated in the study during period I were made aware that they could arrange to be
retroactively paid for their participation. Furthermore, students in period II were now
allowed to specify a time, and a mutually convenient location for data collection. During
period II, no instruments or substantive parts of the methodology were changed. An
additional 20 participants, two groups of 10 participants, agreed to take part in this study.
This brought the total number of participants to equal groups of 20 participants ($N = 40$).

The majority of the participants (22) were enrolled in biology, with eight students
in psychology, seven students in the life sciences and three participants in general
sciences. There were 26 female and 14 male participants who ranged from 18 to 25 years
of age. Table 3 describes the mean academic average, pre-test score, and post-test score
of the 40 participants in this study.

**Creation of multiple-choice test**

Each session began with a test of the prior knowledge of the participants. Students
were asked to answer an 18-question multiple-choice test. The first page of this test asked
students to provide their age, gender, program of study, year of study, prior experience
with any animal physiology courses, prior experience with a specific animal physiology
course and their grade point average in their program. Students were informed that this information would be beneficial for the study, but if they did not wish to they were not obligated to answer any of these questions. These variables were used in the regression analysis as described in the next section.

Each of the 18 test questions had only one correct response, and was designed with the aid of the two biology professors. The test was created with the intention of having a large separation in scores, achieved by creating questions with a wide range of difficulty. To accomplish this one professor selected 18 questions from an exam bank of questions that he used in his past animal physiology exam. He selected six easy questions that were answered correctly by approximately 90% of his students. Then he selected six questions that were moderately difficult that were answered correctly by approximately 60% of his students. Finally, he selected six questions that were difficult typically answered correctly by approximately 30% of his students. The second professor was then asked to modify the questions and to arrange them in ascending order of difficulty.

Furthermore, this test was designed and piloted on biology and physiology students to check that it would not have a ceiling or floor effect when employed. Ideally the median score would be approximately half of the total number of questions, and the distribution would follow a normal curve. Several masters’ students in neuroscience and biology were unable to answer all the questions correctly with the aid of a textbook. After the test was administered to the participants of this study, it was found that no student received a perfect score or a score of zero in the pre-test or post-test.
Multiple-choice test: Animal physiology

The following questions were used as the 18 question multiple-choice pre-test and post-test. The correct answer to each question is highlighted in bold font. These questions were created by the previously mentioned biology professors and taken from their exam bank. They are ordered in the same way as they were presented to the participants.

1. Which of the following are produced by pepsin activity?
   a. Proteoses
   b. Peptones
   c. Peptides
   d. Amino Acids
   e. All of the above

2. Which of the following nitrogenous wastes is the most economical to produce as an excretory strategy?
   a. Nitrogen gas
   b. Ammonia
   c. Urea
   d. Uric acid

3. Which of the following nitrogenous wastes is the most toxic?
   a. Nitrogen gas
   b. Ammonia
   c. Urea
   d. Uric acid

4. The ____________ is the region of the nephron where most of the reabsorption occurs.
   a. Bowman’s capsule
   b. Proximal tubule
   c. Loop of Henle
   d. Distal tubule
   e. Collecting duct
f. Amylase in the mouth initiates the chemical breakdown of starches

5. Which of the following nitrogenous wastes is the least toxic?
   a. Nitrogen gas
   b. Ammonia
   c. Urea
   d. Uric acid

6. Which of the following nitrogenous wastes is the least economical to produce as an excretory strategy?
   a. Nitrogen gas
   b. Ammonia
   c. Urea
   d. Uric acid

7. An acidosis is caused by:
   a. A decrease in $P_{CO_2}$ in the blood
   b. A decrease in carbon dioxide concentration in the blood
   c. Rebreathing
   d. An increase in urea

8. Which treatment(s) cause(s) all digestive enzymes to denature:
   a. Boiling
   b. Freezing
   c. Acidification
   d. All of the above

9. For the most part, freshwater invertebrates and vertebrates are generally
   a. Euryhaline
   b. Hyposomotic
   c. Osmoconformers
   d. Osmoregulators
   e. Ionoconformers

10. For the most part, marine vertebrates are generally
    a. Euryhaline
    b. Hyposomotic
c. Osmoconformers
d. Osmoregulators
e. Ionoconformers

11. Which of the following tends to increase blood glucose?
   a. Cortisol
   b. Epinephrine
   c. Insulin
d. **Glucagon**

12. Which of the following tends to decrease blood glucose?
   a. Cortisol
   b. Epinephrine
c. **Insulin**
d. Glucagon

13. Which of the following statements about glomerular pressure is true?
   a. An increase in glomerular pressure can result from a decrease in the radius of the afferent arteriole.
   b. An increase in glomerular pressure can result from an increase in the radius of the efferent arteriole.
c. **A decrease in glomerular pressure can result from an increase in the radius of the efferent arteriole.**
d. Options A and C are true.

14. Which of the following statements about kidneys is false?
   a. **Though most vertebrates have kidneys, some marine fish have lost their kidneys.**
   b. The kidney tubule is composed of a single cell layer.
   c. Kidney function is augmented by salt glands in marine birds, and rectal glands in sharks.
   d. The loop of Henle allows mammals to make concentrated urine.
e. The renal medulla of beavers is thinner than that of a rabbit because the beaver has shorter loops of Henle.

15. Which animals have the largest ratio of canine teeth to molars?
   a. Mouse
b. Lion  
c. Aardvark  
d. Dolphin

16. The buffering system that is the fastest to respond to pH is:  
a. Electrical  
b. Physical  
c. Chemical

17. Which of the following is NOT a basic property of transport epithelium?  
a. Cellular diversity  
b. Abundant mitochondria  
c. Multiple intercellular connection  
d. Different profiles of proteins in apical and basolateral membranes  
e. **Thick layer of material impervious to ion and water movement.**

18. Which of the following statements about digestive physiology is false?  
a. **Bile salts in combination with pancreatic amylase breaks down dietary lipids**  
b. Cellulase is not produced by animals but symbionts can produce it to aid in the digestion of complex carbohydrate  
c. Pepsin is a proteolytic enzyme that works in the stomach

**Creation of maps**

The participants of this study were charged with the task of studying a concept map or a comparable visual organizer with limited hierarchy, which illustrated the concept of animal physiology. I asked the two biology professors selected for this study to collaboratively create an ideal concept map that contains a rich hierarchy. Initially, the professors decided what propositions they wished to convey regarding the topic of animal physiology. Once a list of propositions was created, the professors were informed of the proper method of creating a concept map. A report by Cañas *et al.* (2003) delivers the
standard method of creating a concept map as described by Novak and other leading concept mapping experts:

The procedure starts with the definition of the topic or focus question. It is critically important to take the time to do this since a Concept Map can lack focus, and the most typical reason why a map might end up in such a state is because of the lack of a clear idea of what the mapper is trying to represent. The goal of the concept mapping session has major impact on the nature of the focus question. If, for instance, the purpose of the Concept Map is to determine what a student knows, then the teacher must have a question clearly in mind. If, on the other hand, the purpose is to elicit knowledge from an expert, the focus question might be more difficult to formulate initially, and may undergo revision as the session goes forward. In the standard Concept Mapping process, the basic steps after identification of the focus question are identifying and listing the most important or “general” concepts that are associated with the topic, ordering the concepts from top to bottom in the mapping field, and adding and labeling linking phrases. Once the preliminary Concept Map has been built, cross-links are identified and added, and a review of the map for completeness and correctness is performed (Cañas et al., 2003, p. 62–63).

The creation of this ideal map was a collaborative process and, as mentioned previously, when students are asked to collaborate on the creation of a concept map, large
groups can defer to the judgment of a selected few (Shen et al., 2007; Roth & Roychoudhury, 1992). This did not occur in this thesis due to the fact that there were only two professors who were asked to collaborate. Also, the professors were asked to participate individually by creating idiosyncratic maps before creating the collaborative map; this exercise was effective in eliciting participation from all members of the group (Shen et al., 2007; Chiu, 2003). As previously mentioned, the visual organizers were created using computer software since it ensures that the maps are legible and consistent for all the participants in the study.

After the concept map was created, the professors were once again asked to individually create another visual organizer conveying the same content, but using limited hierarchy. Once the idiosyncratic maps were completed, the professors collaboratively created the visual organizer with limited hierarchy. Throughout this process, I played a facilitating role in the creation of these maps by providing assistance in the use of software and map design. Both maps were tested again by an independent tester, a university professor with expertise in science, who ensured that the maps contained all the propositions that were previously listed.

**Visual organizers**

The figures below depict the concept map, Figure 1, and the visual organizer with limited hierarchy, Figure 2, used in this study. These visual organizers were created by the professors to answer the 18-question multiple-choice test used to measure the students’ recall of information about animal physiology. The concept map was designed
with the following key features as described by Novak and other concept mapping experts: semi-hierarchical organization, labeled links, and nodes (Cañas et al., 2003).

The specific text used in the visual organizers was selected solely by the two professors who designed them. As previously mentioned, the nodes reflected the anchoring ideas within the concept and the linking phrases specified the nature of the relationship between these ideas. In this regard the visual organizer with limited hierarchy tended to be more descriptive than the concept maps in both the length and specificity of their linking phrases.

The visual organizer with limited hierarchy was created using almost identical text as the concept map; however, it lacked hierarchy both in structure and coloring. The coloring of the concept map and the clustering of the nodes were decided collaboratively by the two professors who designed these maps. They were advised that concept maps are typically designed with fidelity to meaning as opposed to aesthetics or simplicity.

The visual organizer with limited hierarchy contains a node-to-link ratio of exactly one, and can therefore be characterized as a simple visual organizer (Meagher, 2009). The concept map contains almost identical text and nodes, but, has slightly more links than nodes. This concept map would still be thought of as simple despite being slightly more complex than the other visual organizer (Meagher, 2009).

As previously mentioned both the concept map and the visual organizer with limited hierarchy contained all the answers for the 18-questions multiple choice test, this was verified by an independent assessor (a biology professor). These diagrams were
tailored specifically to answer the multiple choice test directly and very few additional nodes (nodes that did not answer a question directly) were added. On the rare occasion that a node on a map did not directly answer a question, then it was put in place to lend clarity to the concept and assist the student in answering the questions correctly. In this regard, all the nodes on the concept map and visual organizer lacking hierarchy were designed to either contribute to the intrinsic load or the germane load of the student. There were no nodes added to intentionally increase the extraneous load of the participants by making questions harder to answer. In this study it is speculated that the concept map may have more germane load than the visual organizer lacking hierarchy because of the hierarchy present in the map. This will be confirmed if the group that uses the concept map is able to recall more propositions about animal physiology than the group that studies the visual organizer with limited hierarchy.
Figure 1: Animal physiology concept map. This concept map explains the concept of animal physiology with the intention of being used to answer the 18-question multiple-choice pre-test and post-test.
Figure 2: Animal physiology visual organizer. This visual organizer with limited hierarchy explains the concept of animal physiology with the intention of being used to answer the 18-question multiple-choice pre-test and post-test.
Data collection

After writing the 18-question multiple-choice pre-test, the students were divided into one of two groups using random assignment. The first group was presented with a concept map that explained the concept of animal physiology. The second group was provided with a visual organizer comparable to the concept map; however, it was limited in its use of hierarchy. Both of these diagrams were magnified and printed on legal size paper (8.5 × 14 inches in size). Each group had precisely 15 minutes to study their visual organizer. This amount of time has been used in other similar concept mapping experiments that require free recall of information (Hall et al., 1999) and provided ample time for students to incorporate this concept into their memory.

The students were then asked to complete a 10-minute poetry exercise. In this time the students were given written instruction on how to write a Haiku. A Haiku is a Japanese poem that is structured to contain three lines that relate to nature. Hiraga (1999) provides a concise description of this art form:

Haiku or hokku as it was called during the time of Basho (1644-1694), is the shortest form of Japanese traditional poetry, consisting of seventeen morae, divided into three sections of 5-7-5. Originating in the first three lines of the 31-mora tanka, haiku began to rival the older form in the Edo period (1603-1867). It was elevated to the level of a profoundly serious art form by the great master Basho. It has since remained the most popular poetic form in Japan. Originally,
the subject matter of haiku was restricted to an objective description of nature suggestive of one of the seasons, evoking a definite, though un-stated, emotional response. Later, its subject range was broadened but it remained the art of expressing as much as possible in the fewest possible words (p. 462).

This Haiku exercise was a distraction that required enough cognitive energy on the part of the participants that they were unable to retain the nodes from the animal physiology exercise in their short-term memory (Meunier & Meunier, 1972). To measure the recall of long-term memory it is essential that one clears the short-term memory of the participants (Schnottz & Kürschner, 2007; Meunier & Meunier, 1972). As a result, students were asked to write haikus about topics far removed from the concept of animal physiology. The students were then asked to re-write the 18-question multiple-choice test which measured their recall of information about animal physiology.

Data analysis

This thesis was quantitative in nature and engaged a variety of statistical tests. To begin with the prior knowledge of the students was gathered and ANOVA tests were used to ensure there were no significant differences in prior knowledge between groups. In comparable studies, prior knowledge was thought of as having two binary categories: high prior knowledge and low prior knowledge. In this study, however, prior knowledge was treated as a continuous variable; this decision was made for several reasons. First, it was speculated that the majority of the participants would be reasonably familiar with the
concept of animal physiology. This topic enters North American curriculum before high
school, so the majority of the students in an undergraduate biology class were expected to
be well acquainted with this topic (Brown, 2003). Subsequently, if one were to separate
low and high prior knowledge groups at the median score then it would be mistakenly
assumed that there are major differences in topic familiarity between these two groups.
This problem would be exacerbated by the limited number of participants in this study (N
= 40).

If the assessment is well designed and does not contain a ceiling or floor effect it
would be reasonable to assume a model of prior knowledge that fits a normal distribution
curve. If this is the case, one could reasonably expect approximately 67% of the
participants to fall in between one standard deviation above and below the median. One
standard deviation difference may not be a substantial difference in the scores attained on
the prior knowledge assessment, especially when one considers the brevity of the test. It
would be dishonest to represent these two categories as high and low prior knowledge
since the majority of the participants would have similar scores. In addition, the data
collected for this thesis underwent tests for normality and no aberrations were found.

A regression analysis was conducted with post-test scores as the outcome
variables; and the age, gender, pre-test scores, year of study, program of study, prior
experience with animal physiology, and academic success of the participants as
predictors. Then a series of set regression analyses was conducted that accounted for
prior knowledge as a continuous variable and sought correlations between the
aforementioned variables. A one-way ANOVA was then conducted to see if there was a significant difference ($p<0.05$) between pre-test and post-test scores on the multiple-choice assessment within each experimental group. Then a two-way ANOVA was conducted that determined if the difference in growth between the groups was significant.

**Alternative analysis**

The pre-test scores of the two experimental groups of students were significantly different [$F(1,38) = 5.86, p = .020$] despite the reasonably large number of participants and their random assignment into their experimental groups. This was thought to be unlikely considering the wide range of difficulty in the questions and the prominent exposure of this topic in the curriculum, however, it did occur. A step-wise regression analysis and repeated measures ANOVA were used to account for this difference in prior knowledge by incorporating the pre-test score in the dependent variable.

**Definition of variables**

The following variables were used in the subsequent analysis: academic average, age, gender, year of study, program of study, pre-test score, post-test score, prior experience with any animal physiology courses, and prior experience with a specific animal physiology course. These variables were described to the participants of the study as follows:

**Academic average.** An estimation by the participant of their current academic average that they have obtained thus far in their current undergraduate degree. For those students who were in their bachelors of education, their cumulative average from their
recent biology undergraduate degree was used. This variable was considered to be on a ratio scale in subsequent analyses and academic averages ranged from 67% to 90%.

**Age.** The current age of the participant. This variable was considered to be on a ratio scale in subsequent analyses and the ages ranged from 18 to 25.

**Gender.** The gender that the participant felt best represents them (a blank space was provided). This variable was considered nominal in subsequent analyses.

**Year of study.** The number of years that the participant had been studying science at the undergraduate level. If the participant had recently graduated from a degree, this variable reflected the number of years that student studied undergraduate science in a post-secondary institution. This variable was considered to be on an interval scale in subsequent analyses and the year of study ranged from 1 to 6.

**Program of study.** The program of study represents their current undergraduate science degree. Four bachelor of education students who had recently completed their undergraduate degree in biology participated in this study and they were considered biology students in the analyses. This variable was considered to be on an ordinal scale in subsequent analyses.

**Pre-test score.** The score obtained by the participant in the 18-question multiple-choice pre-test. There was only one correct answer per question and the maximum score was 18. This variable was considered to be on a ratio scale in subsequent analyses and the scores ranged from 3 to 13.
**Post-test score.** The score obtained by the participant in the 18-question multiple-choice post-test. This test was identical to the pre-test and the participants were made aware during the studying of their visual organizer that would have to write an identical post-test. This variable was considered to be on a ratio scale in subsequent analyses and the scores ranged from 6 to 17.

**Prior experience with any animal physiology courses.** The participants were asked to answer if they had taken any animal physiology courses in the past, and if so, to write the name of such courses. This variable was considered nominal in subsequent analyses.

**Prior experience with a specific animal physiology course.** The participants were asked if they had taken a specific animal physiology course. The name of this course cannot be divulged to protect the anonymity of the professors in the course, but, it should be said that the questions on the assessment and the visual organizers were tailored to the requirements of this one particular animal physiology course. This variable was considered nominal in subsequent analyses.

**Ethics and incentives**

This thesis took into account many ethical considerations. I have a great regard for the autonomy and well-being of the participants. To encourage participation every student in the biology, life science and psychology departments was offered a concept mapping workshop at the end of the study. This workshop provided the results of this study, as well as concept mapping techniques that would optimize student recall of
science topics. This should not be a form of coercion keeping otherwise unwilling participants in the study. Therefore, the workshop was offered for everyone interested regardless of whether or not they agree to be a participant of this study.

Each participant was required to contribute less than 45 minutes of their time to this research project. This is a reasonably small time commitment and should not have caused significant duress to the participants. Also, the topic of this study and the ensuing recall of information were designed to be potentially useful for the participants.

The identity of the participants in this study was protected to the extent possible. The data collected was locked in a storage device when not in use. Furthermore, any archived electronic data was compressed and password-protected. Lastly, the participants’ identities were confidential and reporting of the findings did not present real names of any of the participants or professors. Since the participant pool is extremely large, and the report did not divulge personal details about the population, beyond the average age and gender makeup, the participants’ identities are well protected.

**Summary of methodology**

In this study 40 science students at a mid-sized university in Southern Ontario were recruited and randomly assigned to one of two groups. Both groups were asked to complete an 18-question multiple-choice test on the topic of animal physiology. These participants were then charged with the task of studying one of two visual organizers that illustrated the concept of animal physiology. However, the diagram presented to each group varied in the level of hierarchy used.
After a poetry exercise to flush their memory, these students wrote the same multiple-choice test which measured their recall of information about animal physiology. The data collected from the two groups was then analyzed using regression analyses, ANOVA, and repeated measures ANOVA to determine if there are significant differences in recall depending on the degree of hierarchy present in the instruction. Furthermore, regression analysis, ANCOVA, and repeated measures ANCOVA were used to determine the effect of prior knowledge on their recall.
Chapter 5

Results

Overview of the results

In this section the participants’ data are analyzed using scatter plots for each period of data collection. The data is then amalgamated and a regression analysis and correlational matrix is used to detect and explain statistically significant relationships between the variables. Initially, the pre-test scores between experimental groups were found to be significantly different \[ F(1,38) = 5.86, p = .020 \]. Furthermore, it was found that the post-test scores between the two groups were not significantly different \[ F(1,38) = 0.352, p = .557 \]. After accommodating for the significant difference in pre-test scores between the groups, using ANCOVA, one finds that the difference between the two groups narrows; however, remained non-significant \[ F(1,38) = 2.10, p = .155 \].

Next, the significance of the growth between concept mapping groups and visual organizers lacking hierarchy, from pre-test to post-test scores, was determined using a series of repeated measures ANOVA. A 2-map-type by 2-pre/post test score analysis of variance was conducted with repeated measures on the second factor. It was found that the concept mapping group increased significantly more in their test scores than the visual organizer group \[ F(1,38) = 7.70, p = .009 \].
**Period I: Distribution of pre-test scores and post-test scores**

In the initial period of data collection 20 participants were recruited. The scatter plot below (Figure 3) depicts the pre-test and post-test score of the first 20 participants in this study. The participants using the visual organizer with limited hierarchy are represented by blue dots and the participants using the concept map are represented by green triangles. The $R^2$ of the line of best fit is .029 ($p < .05$). With regards to the participants’ recall of information about animal physiology while using the concept map, one can observe from the scatter plot in Figure 3 that six participants were above this line and four participants below. With regards to the participants using the visual organizer with limited hierarchy, one can observe from the scatter plot that four participants were above this line and six participants below.

In Figure 4, the same points are depicted but they are this time separated by grid lines denoting the average score of the participants. The average pre-test score is 7.85 with a line intersecting the x-axis at this point, and the average post-test score is 13.55 with a line intersecting this point on the y-axis. There are two quadrants that are particularly noteworthy, firstly, the top left quadrant representing participants that had below average pre-test score, but an above average post-test. This quartile contains five participants using the concept maps and one participant using the visual organizer with limited hierarchy. Secondly, the bottom right quadrant which represents the participants that had above average pre-test scores, but below average post-test scores. The bottom right quartile contains four participants that used the visual organizer with limited hierarchy and only one participant who was given a concept map.
Period II: Distribution of pre-test scores and post-test scores

In the second period of data collection 20 more participants were recruited. This time the students were monetarily compensated for their participation, and allowed to decide a convenient time and place to conduct the data collection. The scatter plot below (Figure 5) depicts the pre-test and post-test score of these students. The participants using the visual organizer with limited hierarchy are represented by blue dots and the participants using the concept map are represented by green triangles. The $R^2$ of the line of best fit is .149 ($p < .05$). By observing the trend line and the distribution of participants, one can see that the distribution of scores of the participants in period II do not drastically differ from the distribution of scores of the participants in period I.

Descriptive statistics

Table 3 describes the mean academic average, pre-test score, and post-test score of the 40 participants in this study. There was a significant difference in the pre-test scores between the concept mapping and the visual organizer with limited hierarchy groups [$F(1,38) = 5.86, p = .020$].
Figure 3: Period I distribution scatter plot. This is a scatter plot depicting the pre-test and post-test score of the first 20 participants recruited in this study with a line of best fit ($R^2 = .029$, $p_s < .05$). The participants using the visual organizer with limited hierarchy are represented by blue dots and the participants using the concept map are represented by green triangles.
Figure 4: Period I quartile distribution scatter plot. This is a scatter plot depicting the pre-test and post-test score of the first 20 participants recruited in this study. These participants are separated into quadrants based on the average score of the pre-test (7.85) and post-test (13.55).
Table 3: Descriptive statistics of participants mean pre-test score, post-test score, and academic average

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Average</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concept map</td>
<td>78.83</td>
<td>4.483</td>
</tr>
<tr>
<td>Visual Organizer*</td>
<td>80.20</td>
<td>6.429</td>
</tr>
<tr>
<td>Combined</td>
<td>79.52</td>
<td>5.514</td>
</tr>
<tr>
<td>Pre-test Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concept map</td>
<td>6.85</td>
<td>2.815</td>
</tr>
<tr>
<td>Visual Organizer</td>
<td>8.75</td>
<td>2.099</td>
</tr>
<tr>
<td>Combined</td>
<td>7.80</td>
<td>2.633</td>
</tr>
<tr>
<td>Post-test Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concept map</td>
<td>13.50</td>
<td>1.906</td>
</tr>
<tr>
<td>Visual Organizer</td>
<td>13.10</td>
<td>2.337</td>
</tr>
<tr>
<td>Combined</td>
<td>13.30</td>
<td>2.115</td>
</tr>
</tbody>
</table>

Note: *Visual organizer with limited hierarchy

N = 40

Correlations between variables

The following variables were used in the subsequent analyses: academic average, age, gender, year of study, program of study, pre-test score, post-test score, prior experience with animal physiology courses, and prior experience with a specific animal physiology course. These variables are defined in the Definition of Variables section. The following correlation table was created to determine the dependence of these variables to one another.
Figure 5: Period II distribution scatter plot. This is a scatter plot depicting the pre-test and post-test score of the second group of 20 participants recruited in this study with a line of best fit ($R^2 = .149$, $p < .05$). The participants using the visual organizer with limited hierarchy are represented by blue dots and the participants using the concept map are represented by green triangles.
<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>0.246</td>
<td>0.866</td>
<td>-0.362*</td>
<td>0.375</td>
<td>-0.077</td>
<td>-0.085</td>
<td>-0.216</td>
<td>0.301</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.246</td>
<td>1</td>
<td>0.144</td>
<td>-0.307</td>
<td>-0.426</td>
<td>-0.101</td>
<td>0.032</td>
<td>0.438</td>
<td>-0.142</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.866</td>
<td>0.144</td>
<td>1</td>
<td>0.761**</td>
<td>0.334**</td>
<td>-0.365*</td>
<td>0.343*</td>
<td>0.211</td>
<td>-0.146</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>-0.362*</td>
<td>-0.307</td>
<td>0.761**</td>
<td>1</td>
<td>0.372</td>
<td>0.140</td>
<td>0.058</td>
<td>0.251</td>
<td>-0.111</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>0.375</td>
<td>-0.426</td>
<td>0.334**</td>
<td>0.372</td>
<td>1</td>
<td>0.257</td>
<td>0.256</td>
<td>0.427</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>-0.077</td>
<td>-0.101</td>
<td>-0.365*</td>
<td>0.140</td>
<td>0.257</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>-0.216</td>
<td>0.032</td>
<td>0.343*</td>
<td>0.058</td>
<td>0.256</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>0.301</td>
<td>0.438</td>
<td>0.211</td>
<td>0.251</td>
<td>0.427</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>-0.142</td>
<td>0.438</td>
<td>0.211</td>
<td>-0.111</td>
<td>-0.146</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Note: *Indicates p<0.05
N = 40
In Table 4 there are 13 significant correlations at p>0.05 level of significance, with six of these correlations having p>0.01. Two variables correlated significantly with post-test score, which is the dependent variable of the following regression analyses. It was found that the participants’ post-test scores were correlated to their program of study \( (r = -.376, p = .017) \), and the participants’ prior experience with any animal physiology course \( (r = .385, p = .014) \). Furthermore, the program of study of the participants was correlated with their year of study \( (r = .362, p = .022) \), their prior experience with any animal physiology course \( (r = -.433, p = .005) \), and their prior experience with a specific animal physiology course \( (r = -.521, p = .001) \).

The year of study of the participants was correlated with age \( (r = .681, p < .001) \), their prior experience with any animal physiology course \( (r = .426, p = .006) \), and their prior experience with a specific animal physiology course \( (r = .343, p = .030) \). The participants’ academic average had a direct correlation with their pre-test score \( (r = .372, p = .018) \). Furthermore, the participants’ pre-test score had a positive correlation with their prior experience with any animal physiology course \( (r = .475, p = .002) \), and their prior experience with a specific animal physiology course \( (r = .386, p = .014) \). Finally, it was found that the participants’ prior experience with any animal physiology is correlated to their prior experience with a specific animal physiology course \( (r = .761, p < .001) \).

**Distributions for periods I and II**

The scatter plot below (Figure 6) depicts the pre-test and post-test score of all 40 participants in this study. The participants using the visual organizer with limited
hierarchy are represented by blue dots and the participants using the concept map are represented by green triangles. The $R^2$ of the line of best fit is .091 ($p < .05$). A cluster of five participants in the far left of the x-axis, belonging to the concept mapping group, scored worse on the pre-test than the remaining participants. These five participants were dropped in the final repeated measures ANOVA to mitigate the significant difference in pre-test scores between the experimental groups.

**Regression analyses of individual outcomes in conjunction with map-type predicting post-test scores**

A series of regression analyses were conducted independently testing one variable and the map-type (as a second predictor) in their ability to predict the post-test scores. The following outcomes were used in the analysis: academic average, age, gender, year of study, program of study, pre-test score, prior experience with animal physiology courses, and prior experience with a specific animal physiology course. The last outcome describes whether or not the participants took the course from which these exam bank questions and concept maps were modeled. The results of these tests are reported in Table 5 and the variance is reported in Table 6.
Figure 6: Period I and II distribution scatter plot. This is a scatter plot depicting the pre-test and post-test score of all 40 participants recruited in this study. The participants using the visual organizer with limited hierarchy are represented by blue dots and the participants using the concept map are represented by green triangles.
Table 5: Independent regression analysis of each factor with map-type predicting post-test scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>SE</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic average</td>
<td>.105</td>
<td>.061</td>
<td>.274</td>
</tr>
<tr>
<td>Age</td>
<td>.174</td>
<td>.188</td>
<td>.150</td>
</tr>
<tr>
<td>Gender</td>
<td>1.080</td>
<td>.711</td>
<td>.247</td>
</tr>
<tr>
<td>Year of study</td>
<td>.499</td>
<td>.270</td>
<td>.292</td>
</tr>
<tr>
<td>Program of study</td>
<td>-.782*</td>
<td>.315</td>
<td>-.376</td>
</tr>
<tr>
<td>Pre-test Score</td>
<td>.312*</td>
<td>.132</td>
<td>.388</td>
</tr>
<tr>
<td>Prior experience with any animal physiology course</td>
<td>1.651*</td>
<td>.637</td>
<td>.391</td>
</tr>
<tr>
<td>Prior experience with a specific animal physiology course</td>
<td>1.119</td>
<td>.723</td>
<td>.246</td>
</tr>
</tbody>
</table>

Note: Significance is reported only for b values.
df = 2, 37
*Indicates $p<0.05$

The regression analysis of each individual factor in conjunction with map-type predicting post-test scores yielded three significant results at $p<.05$. The three significant variables are program of study, pre-test score, and prior experience with any animal physiology course with p-values of .018, .023, and .014 respectively. Two other variables should be considered to be viable predictors as well, these are year of study and academic average that had p-values of .072 and .093 respectively. Although, not significant at $p<.05$, without a large N it could be considered a type II error not to report such results.

The variance of the preceding regression analysis is reported in the following table.
Table 6: Variance accounted for by outcomes in conjunction with map type predicting post-test score

<table>
<thead>
<tr>
<th>Variable</th>
<th>$R^2$</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic average</td>
<td>0.083</td>
<td>2.079</td>
</tr>
<tr>
<td>Age</td>
<td>0.032</td>
<td>2.137</td>
</tr>
<tr>
<td>Gender</td>
<td>0.067</td>
<td>2.097</td>
</tr>
<tr>
<td>Year of study</td>
<td>0.093</td>
<td>2.068</td>
</tr>
<tr>
<td>Program of study</td>
<td>0.151</td>
<td>2.001</td>
</tr>
<tr>
<td>Pre-test Score</td>
<td>0.140</td>
<td>2.014</td>
</tr>
<tr>
<td>Prior experience with any animal physiology course</td>
<td>0.162</td>
<td>1.988</td>
</tr>
<tr>
<td>Prior experience with specific animal physiology course</td>
<td>0.069</td>
<td>2.094</td>
</tr>
</tbody>
</table>

*Note: $p_s < .05$, for $R^2$*

The variables that accounted for the most variance are prior experience with any animal physiology course, program of study, and pre-test scores which had $R^2$ values ($p_s < .05$) of .162, .151, and .140 respectively when predicting post-test scores. These three variables combined have an $R^2$ of .250 ($p_s < .05$) with a standard error of 1.933, meaning they account for 25% of the total variance in post-test scores.

**Repeated measures ANOVA**

To determine the significance of the increase in test scores between concept mapping groups and visual organizers with limited hierarchy, from pre-test to post-test scores, a series of repeated measures ANOVA were conducted. A 2-map-type by 2-pre/post test score analysis of variance was conducted with repeated measures on the second factor. The main effect of map-type between subjects was not significant [$F(1,38) = 1.55, p = .221$]; however, the interaction of the map-type and the pre/post test score
was significant \( [F(1,38) = 7.70, p = .009] \). The same analysis was conducted with prior experience with any animal physiology course as a covariate. The interaction of the map-type and the pre/post test score was again significant \( [F(1,38) = 7.39, p = .010] \); however, the interaction of experience with any animal physiology course and the pre/post test score was not significant \( [F(1,38) = 0.802, p = .376] \). Finally, the same test was repeated with the specific animal physiology variable that was previously discussed as the covariate. The interaction of the map-type and the pre/post test score was again significant \( [F(1,38) = 7.78, p = .008] \), and the interaction of prior experience with the specific animal physiology course and the pre/post test score was found to be non-significant \( [F(1,38) = 1.419, p = .241] \).

**Repeated measures ANCOVA accounting for prior knowledge**

In the previous section it was mentioned that a 2-map-type by 2-pre/post test score analysis of variance was conducted with repeated measures on the second factor. When this analysis was run with experience with any animal physiology course as the covariate, the interaction of the map-type and the pre/post test score was again significant \( [F(1,38) = 7.39, p = .010] \); however, the interaction of prior experience with animal physiology and the pre/post test score was not significant \( [F(1,38) = 0.802, p = .376] \). Therefore, prior experience with any animal physiology course does not have a significant effect in increasing the test scores found between the two experimental groups. The same was found for experience with the specific animal physiology course as a covariate, the interaction of the map-type and the pre/post test score was again
significant \[F(1,38) = 7.78, p = .008\], and the interaction of prior experience with the specific animal physiology course and the pre/post test score was found to be non-significant \[F(1,38) = 1.419, p = .241\].

Two other variables should be considered to be viable predictors as well. These are year of study and academic average that had p-values of .072 and .093 respectively. Although, not significant at \( p < .05 \) without a larger N it should be considered a type II error not to report such results. The interaction of the map-type and the pre/post test score was again significant \[F(1,38) = 6.974, p = .012\]; however, the interaction of year of study and the pre/post test score was not significant \[F(1,38) = 0.317, p = .577\]. Finally, the same test was repeated with academic average as the covariate. The interaction of the map-type and the pre/post test score was again significant \[F(1,38) = 7.01, p = .012\], and the interaction of academic average and the pre/post test score was found to be non-significant \[F(1,38) = 0.473, p = .496\].
Chapter 6
Discussion

Overview of the Discussion

This thesis was guided by the following two research questions: Do undergraduate science students using expert-created concept maps differ in their ability to enhance their recall of information about animal physiology when compared to students using visual organizers with limited hierarchy? How does prior knowledge affect the recall of students using concept maps and other visual organizers with limited hierarchy? The data collected from the two experimental groups was analyzed using regression analyses, ANOVA, and repeated-measures ANOVA. It was found that the hierarchical concept-mapping group recalled more of the concept of animal physiology than the visual-organizer group \[F(1,38) = 7.70, \ p = .009\]. This significant difference in growth remains even if one accounts for prior knowledge by inputting the two animal physiology experience variables (independently) as covariates in a repeated-measures ANCOVA. This finding supports the contention that the hierarchical concept map provided a benefit in the recall of information when compared to the visual organizer with limited hierarchy. The additional hierarchy found in the concept maps is thought to elicit germane load in the students. Finally, participants who achieved low scores on the pre-test were removed
and the repeated-measures ANOVA was repeated to mitigate the significant difference in pre-test scores between the experimental groups.

**Period I: Distribution of pre-test scores and post-test scores**

In the first period of recruitment only 20 participants were recruited despite expanding this period from one week to an entire month. During period I, every participant recalled more information about animal physiology in their post-test compared to their pre-test. This occurred regardless of whether or not their visual organizer contained hierarchy; however, there was a non-significant difference between the post-test scores of these two groups \(F(1,18) = 0.435, p = .518\).

This difference is most noticeable in Figure 4, a scatter plot that depicts the pre-test and post-test score of every participant in period I \((N = 20)\). In this figure, the participants are depicted with grid lines denoting the average score of the participants. The average pre-test score is 7.85 with a line intersecting the x-axis at this point, and the average post-test score is 13.55 with a line intersecting this point on the y-axis.

There are two quadrants that are particularly noteworthy, firstly, the top left quadrant representing participants that had below average pre-test score, but an above average post-test. This quartile contains five participants using the concept maps and one participant using the visual organizer with limited hierarchy. Secondly, the bottom right quadrant which represents the participants that had above average pre-test scores, but below average post-test scores. The bottom right quartile contains four participants that
used the visual organizer with limited hierarchy and only one participant who was given a concept map.

Unfortunately, the number of participants was too small to draw any reasonable conclusions. Ultimately, these trends are entirely inferential and have no statistical rigor. The most that can be drawn from these observations is encouragement that this trend may continue when the number of participants in this study increases. This initial analysis lead to the initiation of the second period of data collection doubling the total participant pool (N=40).

**Correlations between predictors**

As previously mentioned, and can be seen by Table 4, there are 13 significant correlations at \( p < .05 \) level of significance, with six of these correlations having \( p < .01 \). Two variables correlated significantly with post-test score, which is the dependent variable of the following regression analyses. It was found that the participants’ post-test scores were correlated to their program of study (\( r = -.376, p = .017 \)), and the participants’ prior experience with any animal physiology course (\( r = .385, p = .014 \)). The first correlation shows that participants whose programs of study included more biology courses tended to perform better on the post test. Both of these results support a well-designed post-test since, as predicted, students who had previous experience with animal physiology and biology courses tended to score better. Of further interest is a non-statistically significant relationship between the specific animal physiology course and the post-test scores (\( r = .246, p = .127 \)). These correlations seem to indicate that students
did not remember the answer to specific questions that they had answered in the past. However, students who had prior experience with animal physiology had a familiarity with the concept that would allow them to recall this information with greater ease in the future (Sweller & Chandler 1994; Sweller et al., 1998).

The program of study of the participants was related to their prior experience with any animal physiology course \((r = -0.433, p = 0.005)\), and their prior experience with a specific animal physiology course \((r = -0.521, p = 0.001)\). These correlations show that students who were in biology and life science programs tended to have more experience with animal physiology courses than students in psychology and general sciences. The correlation occurs because a physiology course (either animal or plant physiology) is necessary for the completion of a biology degree, in the university where these data were collected. However, this course is an elective for psychology and general science students, and therefore, prior enrollment in this class was less probable.

The year of study of the participants was correlated with age \((r = 0.681, p < 0.001)\), their prior experience with any animal physiology course \((r = 0.426, p = 0.006)\), and their prior experience with a specific animal physiology course \((r = 0.343, p = 0.030)\). Intuitively, students who are closer to finishing their undergraduate degree tended to be older and would also have a greater chance to being exposed to animal physiology courses.

However, age was not significantly correlated with post-test \((r = 0.146, p = 0.370)\) or pre-test scores \((r = -0.025, p = 0.880)\).
The participants’ academic average had a direct correlation with their pre-test score \((r = .372, p = .018)\). Furthermore, the participants pre-test score had a direct correlation with their prior experience with any animal physiology course \((r = .475, p = .002)\), and their prior experience with a specific animal physiology course \((r = .386, p = .014)\). These results are expected when one considers animal physiology courses provided an opportunity for recall prior to the intervention (Schnitz & Kurzchner, 2007). These three correlations further support the legitimacy of the pre-test as a measure of prior knowledge of animal physiology. These findings correspond with other studies that claim students with low prior knowledge reported exerting more mental effort in answering questions than did a high prior knowledge group (Amadieu et al., 2009).

Finally, it was found that the participants’ prior experience with any animal physiology course is correlated to their prior experience with a specific animal physiology course \((r = .761, p < .001)\). The strong significance of this correlation occurs because students who are enrolled in the specific animal physiology course would have previous experience with animal physiology. This correlation indicates that 71\% of the students, who did take animal physiology in the past, took the specific animal physiology course from which the instruments of this study are modeled.

**The benefit of hierarchy**

In this study most of the participants had an increase between the pre-test scores and post-test scores. The two participants who did not belonged to the group using the visual organizer with limited hierarchy. The null hypothesis that there is no difference
between the post-test scores of the two experimental groups could not be rejected 
\[ F(1,38) = 0.352, p = .557 \]. As previously mentioned, the pre-test scores between these 
groups was significantly different \[ F(1,38) = 5.86, p = .020 \]. After accommodating for 
the significant difference in pre-test scores between the groups, using ANCOVA with the 
pre-test scores as a covariate, we find that the difference between the groups is still non-
significant \[ F(1,38) = 2.10, p = .155 \] with a partial $\eta^2 = .054$. With only 40 participants in 
this study and $p = .155$, it is possible that this difference in post-test scores would be 
significant at $p < .05$ had the total number of participants been larger; however, it is 
unlikely that this effect is pedagogically significant due to the small partial $\eta^2$.

The participants in the concept mapping group improved on their post-test score 
on average 6.65 points when compared to their pre-test score. The participants in the 
visual organizer with limited hierarchy group changed on average 4.35 points in the same 
comparison. The concept mapping group had 12.8% greater recall of information about 
animal physiology (measured as a difference between pre-test and post-test scores) than 
the visual organizer group.

This evidence supports the idea that the semi-hierarchical organization of the 
concept map confers some benefit in the recall of information about animal physiology 
(Schnotz & Kürschner, 2007); however, previous analysis of just the post-test scores 
deemed the difference to be non-significant.
To determine the significance of the growth between concept mapping groups and visual organizers with limited hierarchy, from pre-test to post-test scores, a series of repeated measures ANOVA were conducted. A 2-map-type by 2-pre/post test score analysis of variance was conducted with repeated measures on the second factor. From the results of this analysis we can conclude that pre-post test scores, as one factor, is significantly different between the concept mapping group and the visual organizer with limited hierarchy group \([F(1,38) = 7.70, p = .009]\). This significant difference in growth remains even if one accounts for prior knowledge by inputting the two animal physiology experience variables (independently) as covariates. From this we can conclude that the growth in test scores of the students in the concept mapping group, as measured by the 18-question multiple-choice test, is significantly different than the visual organizer group. This finding agrees with other literature discussing concept mapping interventions which speaks favorably of similar interventions providing benefits in the recall of science concepts (Rewey et al., 1989; Mayer & Moreno, 2003; Ruiz-Primo & Shavelson, 1996; Royer & Royer, 2004).

Unfortunately, since the pre-test score is significantly lower in the concept mapping group we cannot assume that this increase in growth would be conferred if the two groups had similar pre-test scores. It is possible that the visual organizers would both produce improvements in the post-test to scores of approximately 13.3. To mitigate this pre-test bias another repeated measure ANOVA excluding participants with the lowest
pre-test scores was conducted which is addressed in section 6.6 Controlling for the significant difference in pre-test scores.

**Effects of prior knowledge**

The variables that accounted for the most variance, when predicting post-test scores in conjunction with map-type, are prior experience with any animal physiology course, program of study, and pre-test scores which had $R^2$ values ($p < .05$) of .162, .151, and .140 respectively. These three variables combined have an $R^2$ of .250 ($p < .05$) with a standard error of 1.933, meaning they account for 25% of the total variance in post-test scores. The program of study was also analyzed as a categorical variable, with life science and biology as a high prior knowledge category and general science and psychology as a low prior knowledge category. This analysis yielded no significant results.

All three of these variables seem to be relevant measures of prior knowledge and they explain a substantial portion of the variance in the post-test scores.

A 2-map-type by 2-pre/post test score analysis of variance was conducted with repeated measures on the second factor and both animal physiology variables independently inputted as covariates. These findings show that regardless of whether or not we control for experience with animal physiology (both variables), academic average, or year of study, the recall of information about animal physiology was found to be significantly different between the two experimental groups ($p < .05$). This is encouraging and further supports the contention that the concept maps provide a benefit
in the recall of information about animal physiology, even when prior knowledge is accounted for between groups (Amadieu et al., 2009). However, the issue of a statistically significant difference in pre-test scores has not yet been addressed, and this is done in the next section.

When viewing a concept map it is assumed that general details would be closer to the topic node with context specific details located farther away from the head node (Cañas et al., 2003). Although both maps have almost identical text, number of nodes and number of links; it is clear that the concept map is better organized as a result of its hierarchy. However, students in the science program may not have extensive experience with concept maps; therefore, when reading the map they are expending greater cognitive effort than those who are adept with concept maps (Amadieu et al., 2009). In similar future studies, it would be beneficial to ascertain the knowledge and level of experience of the participants with regards to concept maps and other visual organizers, an issue discussed later in this thesis (see section Limitations).

**Controlling for the significant difference in pre-test scores**

The significant difference in growth found in the concept mapping group, resembles the aforementioned study by Guastello, Beasley, and Sinatra (2000). In this study, students with low prior-knowledge benefited from a concept mapping intervention. The authors of this study believed that the concept mapping treatment was responsible for a six standard deviation increase in learning when compared to a traditional lecture. However, when this intervention is compared to other similar concept mapping studies
(55 concept mapping interventions) it is seen as a significant outlier (Nesbit & Adesope, 2006).

The exceptionally large difference in standard deviation favoring the experimental group is likely caused by the control group attempting to learn a rather complicated scientific concept without any support from images or diagrams. This issue may be exacerbated if the final assessment demands or encourages students to answer questions with a diagram. Instead of the purported advantage in learning, students in the experimental group may just be provided with ideal answers for the test, an advantage not shared by the treatment group. Unfortunately, the authors of the study did not provide sufficient details about the assessment to properly resolve this issue (Guastello et al., 2000).

Unlike the study by Guastello, Beasley, and Sinatra (2000) this study provides two experimental groups both given visual organizers. The significant difference in recall between these two groups exists despite both groups of participants having considerable prior knowledge and academic success. This is encouraging since differences in recall would most likely be exacerbated in low prior knowledge groups. Finally, this thesis studies the role that low prior knowledge has on the growth of the participants’ recall. Since the pre-test score is significantly lower in the concept mapping group one cannot assume that this increase in growth seen earlier would be conferred if the two groups had similar pre-test scores. As previously mentioned, it is possible that the visual organizers would both produced improvements in the post-test to scores between 13.1 and 13.5. As
seen in Figure 6, there are two participants in the concept mapping group who received a score of three on their pre-test. These participants received the lowest score on the pre-test of all the participants in this study.

After excluding these two participants from the analysis, it was found that the difference in pre-test scores between the two groups was no longer significant \( F(1,36) = 3.67, p = .063 \). Another 2-map-type by 2-pre/post test score analysis of variance was conducted with repeated measures on the second factor. The main effect of map-type between subjects was not significant \( F(1,36) = .755, p = .391 \); however, the interaction of the map-type and the pre/post test score was significant \( F(1,36) = 5.25, p = .028 \). It can be seen from these results that when these two participants are removed there is a non-significant difference in pretest scores, and a significant difference in the pre/post score interaction with map-type. These results lend even greater support to the contention that the concept mapping intervention is more beneficial than the visual organizer intervention. However, one could argue that this conclusion is premature since it is possible to view \( p = 0.63 \) as close enough to \( p < .05 \), that further steps must be taken in the analysis to confirm this conclusion.

To ameliorate this concern, the same analysis was completed excluding the five lowest scoring participants of the pre-test. This criterion is reasonable since it removes all the participants in the concept mapping intervention that scored below the lowest scoring participant in the visual organizer with limited hierarchy group. It was found that the difference in pre-test scores was, as expected, non-significant \( F(1,33) = 1.16, p = .289 \).
The interaction of the map-type and the pre/post test score was also non-significant [F(1,33) = 2.99, \( p = .093 \)]. This result does seem to contradict the preponderance of previous evidence supporting the claim that the concept mapping intervention provided significantly greater recall of information about animal physiology than the visual organizer group. This last result may not be contradictory since the \( p \)-value (\( p < .010 \)) is still reasonably close to \( p < 0.05 \). Furthermore, this last result may be less valid because it has a substantially lower number of participants than the previous analyses (between 7.5 - 12.5%) fewer participants.

Unlike the study by Guastello, Beasley, and Sinatra (2000), it would seem that the advantage in recall provided by concept maps can be extended beyond low prior knowledge students. When two participants are removed to discover the role of prior knowledge on the advantage conferred by concept maps with regards to recall, it was found that the growth is significant. When five participants are removed this trend still remains, albeit no longer significant at \( p > .05 \), which could be a result of the low number of participants (\( N = 35 \)). Using the conceptual framework of modern cognitive load theory (Schnotz & Kürschner, 2007), the additional cognitive load added by the hierarchy in the concept maps appears to be germane load.

**Cognitive load in concept maps**

The purpose of this thesis is to determine the efficacy of semi-hierarchical organization within concept maps in enhancing the recall of information about animal physiology. The concept map provided benefits in recalling information about animal
physiology when compared to the visual organizer with limited hierarchy. It is speculated that the participants using the concept map experienced more cognitive load than the students using the visual organizer lacking hierarchy. Most types of visual organizers are effective at reducing some of the cognitive burden (extraneous load) of recalling a science concept (Cooper et al., 2001). The semi-hierarchical organization of concept maps enhances this benefit. Hierarchy in concept maps is an additional constraint which seems to limit their use as an instructional tool (Guastello et al., 2000). Hierarchy also adds to the overall cognitive load of students; however, it is beneficial and therefore appears to constitute germane load.

As previously mentioned, Ruiz-Primo and Shavelson (1996) declare that the use of concept maps may theoretically increase the extraneous load rather than the germane load, specifically when the instruction is created by an expert instead of the student. Concept maps can be idiosyncratic and if the student did not create the concept map then understanding the meaning of the diagram may add extraneous load. This claim was unsubstantiated in this study since the concept maps, despite being more complex, provided a significant increase in the growth of scores when compared to the visual organizer with limited hierarchy. This coincides with the literature that claims that although there is theoretical merit for extraneous load increasing when learning with concept maps; it is not supported by the results of concept mapping experiments to be a detriment to recall (Nesbit & Adesope, 2006).
Another possibility is that concept maps, in addition to having more germane load, may also have less extraneous load when compared to the visual organizer with limited hierarchy. Extraneous load in visual organizers may emanate from their lack of hierarchy causing nodes to be improperly aligned causing split-attention effect (Mayer & Moreno, 2003). Since hierarchy and linking words between nodes are required in a concept map; text becomes naturally aligned with images and other structures within the map and this mitigates split-attention (Mayer & Moreno, 2003). However, it is unlikely that the concept map in this experiment had less extraneous load than the visual organizer lacking hierarchy since both diagrams used nearly identical text. The concept map in this thesis was built with more complexity than the other visual organizer; as a result it should theoretically contribute more germane load (Ruiz-Primo et al., 2001; Amadieu et al., 2009). This issue could be better resolved if each participant was given a self-report survey (Amadieu et al., 2009) which measured their overall cognitive load during the intervention (see section Future impact of this study).
Chapter 7

Conclusion

The purpose of this thesis is to determine the efficacy of semi-hierarchical organization within concept maps in enhancing recall using the conceptual framework of cognitive load theory. This research is valuable because it fills a gap in the current concept mapping literature by focusing on a controlled experiment that methodically tests a specific characteristic of concept maps (Katayama & Robinson, 2000; Nesbit & Adesope, 2006). This study sought to understand if the constraints of concept maps are meaningful or an unnecessary encumbrance when compared to other visual organizers.

Initially, the pre-test scores between experimental groups were found to be significantly different and this complicated the preliminary findings [F(1,38) = 5.86, p = .020]. Furthermore, it was found that the post-test scores between the two groups were not significantly different [F(1,38) = 0.352, p = .557]. A 2-map-type by 2-pre/post test score analysis of variance was conducted with repeated measures on the second factor. It was found that the concept mapping group increased significantly more in their test scores than the visual organizer group [F(1,38) = 7.70, p = .009]. This finding supports the contention that the use of a concept map provided a greater benefit in the recall of information about animal physiology than a comparable visual organizer with limited hierarchy. The concept mapping group always showed a significantly greater advantage in the growth of test scores when compared to the visual organizer group during repeated-
measure ANCOVA, even when using prior knowledge variables as a covariate. The majority of the research findings in this thesis support the contention that concept maps confer a slight advantage in the recall of information about animal physiology when compared to a comparable visual organizer with limited hierarchy. The semi-hierarchical organization of concept maps are thought to be germane in this experiment based on modern cognitive load theory (Schnitz & Kürschner, 2007; Ruiz-Primo et al., 2001; Amadieu et al., 2009).

**Limitations**

There are a few limitations to this study that may affect the validity of the research findings. One major limitation of this study is the number of participants. Unfortunately, I was unable to find a large number of enthusiastic students who wished to participate in this project. Ideally, I should have had 60 participants or more, which, if achieved, should have been enough to provide much greater measures of significance for most forms of parametric statistical analyses. I was unable to reach the desired number of participants, and this study finished with 40 participants. Furthermore, the desired number of participants is smaller than well-funded national research projects. However, facilitating an increased number of participants would require resources that are currently not at my disposal.

In similar future studies it would also be beneficial to appropriate the knowledge and level of experience of the participants with regard to concept maps and other visual organizers. This would allow me to analyze the prior knowledge of the participants’ more
effectively. In previous studies it was found that participants who had prior experience with concept maps were considerably better at reading them and utilizing them to achieve learning outcomes (Amadieu et al., 2009). If students had experience reading visual organizers this would reduce their extraneous load when reading complex concept maps and visual organizers (Amadieu et al., 2009; Schnitz & Kürschner, 2007).

Another limitation of this study lies in the generalizability of these findings. This research was conducted with a specific type of participant, university students in an undergraduate science program. This group would not necessarily reflect the age, gender ratio, or socioeconomic status of the average Canadian university classroom. To recruit such a large number of participants I would have to draw students from a few different science faculties. This would mean that students would be in various years of their program and have taken varying amounts of physiology courses. Therefore, some students would have a greater knowledge of animal physiology than others. This is not a major concern, since animal physiology enters North American curriculum before high school, so the majority of the students in an undergraduate biology class should be acquainted with this topic (Brown, 2003). Also, the participants were assigned to their experimental groups randomly. Random assignment of participants into their experimental groups did alleviate some of the systematic bias. Regardless, a pre-test was conducted to determine if there was any significant difference in prior knowledge between the experimental groups. Despite the aforementioned factors, a significant difference was found [F(1,38) = 5.86, p = .020]. This difference was accounted for by
using a step-wise regression analysis in post-test comparison. Ultimately, it would have been better if this statistically significant difference had not occurred, which could have been brought about by having experimental groups larger than 20 participants.

If I were to repeat this study, I would improve the efficacy of these results by recruiting more participants. By increasing the number of participants I would be considerably less likely to have the significant difference in pre-test scores that I found. Although, I had expended great effort in recruiting as many students as possible, I had limited myself to one university and I would hope to, in the future, draw participants from many different post-secondary institutions. Secondly, I would attempt to obtain funding so I could pay the participants for their involvement in their study, as I did in the second period of this study. Thirdly, although consistency in the environment mitigates environmental bias; I may have to concede this advantage, as I did in the second period of this study, for the sake of pragmatism. By having sessions conducted throughout the day, and in various rooms, I would be able to provide greater convenience for all of the participants, thereby increasing the recruitment numbers.

**Future impact of this study**

This study has laid the ground work for a doctoral study with 200 participants separated into four experimental groups \((n = 50)\), with students separated by high and low prior knowledge, and the concept map and visual organizers with limited hierarchy groups. By separating the students into high and low prior knowledge groups, one would be able to create a 2x2 ANOVA that would be able to test the effect of prior knowledge
on the map-type with greater specificity. Furthermore, this separation would represent a larger gap in prior knowledge between participants than is present in this current study, and this would increase the likelihood of a prior knowledge affect in the analysis.

This study would also include an exercise that would measure the participants’ cognitive load and attenuation during the study. This would be beneficial in this experiment since it would allow me to confirm the difference in cognitive load that is thought to be occurring. Studies have shown an increase in cognitive load when students are looking at more complex concept maps and visual organizers (Amadieu et al., 2009; Ruiz-Primo et al., 2001; Schnotz & Kürschner, 2007). Increasing the hierarchy present in a concept map will also increase the cognitive load present (Amadieu et al., 2009). A self-report instrument could be administered shortly after the intervention to test the participants’ cognitive load and attenuation during this experiment which would provide further evidence of an increase in cognitive load that is speculated to be occurring.

A larger number of participants would better ensure that no pre-test difference existed between the groups. Also, this study would test students’ prior knowledge and experience with concept maps and other visual organizers. Furthermore, this study would employ a longitudinal design to better understand the duration that the knowledge acquired is retained in long term memory. Finally, with four experimental groups and the self-report instrument (Amadieu 2009), one would be able to better analyze the role of prior knowledge as it would be accurately represented as a categorical variable. This
would also allow one to observe with greater specificity the cognitive load of the participants and this would enhance the results of the thesis.
References


Figure 7. A concept map that describes the literature that was used to inform the methodology of this study.
Appendix B

Anatomy of Concept Maps

Figure 8. A concept map that illustrates the key structural components of concept maps including nodes, linking words, and propositions.
Anatomy of Visual Organizer with limited hierarchy

Figure 9. A graphic that illustrates the key structural components of the visual organizer with limited hierarchy.
Appendix C

Consent Form

A STUDY OF SEMI-HIERARCHICAL ORGANIZATION IN THE CONSTRUCTION OF CONCEPT MAPS USING THE FRAMEWORK OF COGNITIVE LOAD THEORY

Please sign one copy of this Consent Form and return to Dev Thain. Retain the second copy for your records.

Name (please print clearly): ____________________________

1. I have read and retained a copy of the Letter of Information and Consent Form and had questions answered to my satisfaction.

2. I understand that I will be participating in the study called “A STUDY OF SEMI-HIERARCHICAL ORGANIZATION IN THE CONSTRUCTION OF CONCEPT MAPS USING THE FRAMEWORK OF COGNITIVE LOAD THEORY”. The purpose of this research is to examine the benefits of semi-hierarchical organization within concept mapping. The study will require me to attend a single one hour group session, in a classroom in the biosciences building, where I will be asked to take an 18-question multiple-choice test about animal physiology. I will be asked to provide my age, gender, year of study and grade point average. I will then study a visual organizer that contains information about animal physiology. I will be asked to engage in an exercise to create poems, this will not be tested in any way, and instead it will be used to flush my memory stores. Finally, I will complete the same multiple-choice test that was previously shown to me. There are no known physical, psychological, economic, or social risks associated with this study. The total time required by me should not exceed one hour.

3. I understand that my participation in this study is voluntary and I may withdraw at any time without any effect on my standing in school. Should I wish to withdraw I can leave the study at any time without consequence. I may simply leave the data collection session and the data will be destroyed. If I wish to withdraw at a later point, I can email that request. However, if my data has already been submitted, due to the anonymity of the collected data it may not be possible to separate out my data. I understand that every effort will be made to maintain the confidentiality my data to the greatest extent possible. Only the experimenter and faculty advisors will have access to my data. The data may also be published in professional journals or presented at scientific conferences, but any such presentations will be of general findings and will never breach individual confidentiality.

4. Any questions about study participation may be directed to Dev Thain at 4dt3@queensu.ca; or my supervisor, Dr. Reeve at 613-533- 6000 x77296; reever@queensu.ca. Any ethical concerns about the
study may be directed to the Chair of the General Research Ethics Board at 613-533-6081 or chair.GREB@queensu.ca

I have read the above statements and freely consent to participate in this research:

Signature: ___________________________ Date: ____________________

Should you be interested, you are entitled to a copy of the findings, please provide your email if you wish a copy of the results: _____________.

106
Appendix D

Letter of Information

A STUDY OF SEMI-HIERARCHICAL ORGANIZATION IN THE CONSTRUCTION OF CONCEPT MAPS USING THE FRAMEWORK OF COGNITIVE LOAD THEORY

This research is being conducted by Dev Thain under the supervision of Richard Reeve, in the Faculty of Education at Queen’s University in Kingston, Ontario. This study was granted clearance by the General Research Ethics Board for compliance with the Tri-Council Policy Statement: Ethical Conduct of Research Involving Humans, and Queen’s policies.

What is this study about? The purpose of this research is to examine the benefits of semi-hierarchical organization within concept mapping. The study will require participants to attend a single one hour group session, in a classroom in the biosciences building, where they will be asked to take an 18-question multiple-choice test about animal physiology. Students will be asked to provide their age, gender, year of study and grade point average. They will then study a visual organizer that contains information about animal physiology. Participants will be asked to engage in an exercise to create poems, this will not be tested in any way, and instead it will be used to flush the memory stores of the learner. Finally, the participants will complete the same multiple-choice test that was previously shown. There are no known physical, psychological, economic, or social risks associated with this study. The total time required by the participants should not exceed one hour.

Is my participation voluntary? Yes. Participants are not obliged to answer any objectionable or discomforting questions. You may also withdraw at any time with no effect on your standing in school. Should someone wish to withdraw they can leave the study at any time without consequence. The participant may simply leave the data collection session and their data will be destroyed. If the participant wishes to withdraw at a later point, s/he can email me with that request. However, if the participant’s data has already been submitted, due to the anonymity of the collected data it may not be possible to separate out the participant’s data.

What will happen to my responses? We will keep your responses confidential to the greatest extent possible. Only the experimenter and faculty advisors will have access to this data. To help us ensure confidentiality, please do not put your name on any of the research study answer sheets. The data may also be published in professional journals or presented at scientific conferences, but any such presentations will be of general findings and will never breach individual confidentiality. Should you be interested, you are entitled to a copy of the findings, please email me at 4dt3@queensu.ca and the results will be provided to you. In accordance with Queen’s policy, data will be retained for a minimum of five years. This data will not be used for secondary purposes.

Will I be compensated for my participation? No, you will receive no compensation for your time. There will be no remuneration for your participation.

What if I have concerns? Any questions about study participation may be directed to Dev Thain at 4dt3@queensu.ca or my supervisor Dr. Reeve at 613-533-6000 x77296; reever@queensu.ca. Any ethical
concerns about the study may be directed to the Chair of the General Research Ethics Board at chair.GREB@queensu.ca or 613-533-6081.

This study has been granted clearance according to the recommended principles of Canadian ethics guidelines, and Queen's policies.