Using Concept Maps to Monitor Knowledge Structure Changes in a Science Classroom

Leah J. Cook
Western Michigan University, leahjcook@gmail.com

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Part of the Educational Assessment, Evaluation, and Research Commons, and the Science and Mathematics Education Commons

Recommended Citation

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
USING CONCEPT MAPS TO MONITOR KNOWLEDGE STRUCTURE
CHANGES IN A SCIENCE CLASSROOM

by

Leah J. Cook

A dissertation submitted to the Graduate College
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
Science Education
Western Michigan University
June 2017

Doctoral Committee:

Brandy Pleasants, Ph.D., Chair
Amy Bentz, Ph.D.
Chris Hamstra, Ph.D.
The aim of this research is to determine what differences may exist in students’ structural knowledge while using a variety of concept mapping assessments. A concept map can be used as an assessment which connects concepts in a knowledge domain. A single assessment may not be powerful enough to establish how students’ new knowledge relates to prior knowledge. More research is needed to establish how various aspects of the concept mapping task influence the output of map creation by students. Using multiple concept maps and pre-instruction and post-instruction VNOS instruments during a 16-week semester, this study was designed to investigate the impact of concept map training and the impact of assessment design on the created maps. Also, this study was designed to determine what differences can be observed between expert and novice maps and if similarities and differences exist between concept maps and an open-ended assessment. Participants created individual maps and the maps were analyzed for structural complexity, overall structure, and content. The concept maps were then compared by their timing, design, and scores.

The results indicate that concept mapping training does significantly impact the shape and structure complexity of the map created by students. Additionally, these data support that students should be frequently reminded of appropriate concept mapping skills and opportunities
so that good mapping skills will be utilized. Changing the assessment design does appear to be able to impact the overall structure and complexity of created maps, while narrowing the content focus of the map does not necessarily restrict the overall structure or the complexity. Furthermore, significant differences in structural complexity were observed between novice and expert mappers. The fluctuations of NOS concepts identified in student created maps may suggest why some students were still confused or had incorrect conceptions of NOS, despite explicit and reflective instruction throughout the semester.
ACKNOWLEDGEMENTS

I am forever thankful for the patience and guidance provided by Dr. Brandy Pleasants in completing my dissertation. I am thankful for her time and encouragement during this process. I owe similar gratitude to Dr. Amy Bentz and Dr. Chris Hamstra for their encouragement and suggestions. I am grateful to Dr. Andria Bierema and Dr. Renee’ Schwartz for allowing me to collect data and Dr. Sarah Krajewski for her help. I would also like to thank the Mallinison Institute for Science Education in their support of my research and teaching at Western Michigan University. Without all the aforementioned people, my dissertation research would not have been completed.

Finally, my completed dissertation would not have been possible without acknowledging the three pillars in my life: God, my husband, and three children, Charlotte, Lukas, & Joey. Isaiah 43:2, “When you pass through the waters, I will be with you; And through the rivers, they will not overflow you. When you walk through the fire, you will not be scorched, nor will the flame burn you.” DeWayne, thanks for encouraging me to keep pursuing new opportunities and helping and pushing me to finish this goal. I love you and am thankful to you for everything. Charlotte, Lukas, and Joey, you are the legacy of all this work. “Mommy’s presentations and paper” are done. While this is being an earthly goal, may the testimonies of this time growing in my faith resound and encourage others around me. God Bless.

Leah J. Cook
# TABLE OF CONTENTS

**ACKNOWLEDGEMENTS** .................................................................................................................. ii

**LIST OF TABLES** ...................................................................................................................... viii

**LIST OF FIGURES** .................................................................................................................... ix

**CHAPTER I: INTRODUCTION** ................................................................................................. 1  
  - Background .......................................................................................................................... 1
  - Statement of the Problem .................................................................................................. 2
  - Purpose of the Study ......................................................................................................... 4
  - Significance of the Study .................................................................................................. 4
  - Theoretical Framework ..................................................................................................... 5

**CHAPTER II: REVIEW OF THE LITERATURE** ................................................................. 7  
  - Structural Knowledge ..................................................................................................... 7
    - Knowledge Structure Construction ........................................................................... 8
    - Structural Knowledge Map Contributions ............................................................... 10
    - Structural Knowledge Map Limitations ................................................................... 11
    - Summary .................................................................................................................... 12
  - Assessment in a Classroom ............................................................................................... 13
    - Historical Context of Assessment ............................................................................ 13
    - Types of Assessment .................................................................................................. 16
    - Teachers Impacting Assessment ............................................................................... 19
    - Assessment Impacting Students ............................................................................... 20
    - Influences of Assessment ......................................................................................... 21
Table of Contents - Continued

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment Limitations</td>
<td>22</td>
</tr>
<tr>
<td>Conclusions and Research Gaps</td>
<td>24</td>
</tr>
<tr>
<td>Summary</td>
<td>24</td>
</tr>
<tr>
<td>Concept Mapping Knowledge Structures as Assessment</td>
<td>25</td>
</tr>
<tr>
<td>History of Concept Mapping</td>
<td>26</td>
</tr>
<tr>
<td>Concept Mapping as Summative Assessment</td>
<td>27</td>
</tr>
<tr>
<td>Concept Mapping as Formative Assessment</td>
<td>28</td>
</tr>
<tr>
<td>Major Research Concept Mapping Themes</td>
<td>30</td>
</tr>
<tr>
<td>Concept Mapping with other Assessments</td>
<td>30</td>
</tr>
<tr>
<td>Expert versus Novice Aspects of Concept Mapping</td>
<td>32</td>
</tr>
<tr>
<td>Summary</td>
<td>34</td>
</tr>
<tr>
<td>Research Questions</td>
<td>35</td>
</tr>
<tr>
<td>CHAPTER III: METHODOLOGY</td>
<td>37</td>
</tr>
<tr>
<td>Class and Sample Description</td>
<td>37</td>
</tr>
<tr>
<td>Data Sources and Timeline</td>
<td>37</td>
</tr>
<tr>
<td>Concept Maps</td>
<td>38</td>
</tr>
<tr>
<td>VNOS Instrument</td>
<td>43</td>
</tr>
<tr>
<td>Interviews</td>
<td>43</td>
</tr>
<tr>
<td>Data Analyses</td>
<td>43</td>
</tr>
<tr>
<td>Concept Map Analysis</td>
<td>43</td>
</tr>
<tr>
<td>Statistical Analyses</td>
<td>46</td>
</tr>
<tr>
<td>VNOS Analysis</td>
<td>47</td>
</tr>
</tbody>
</table>
Table of Contents - Continued

Confidentiality .......................................................................................................................... 47
Role of the Researcher ............................................................................................................ 47
Limitations to the Study ......................................................................................................... 48
Summary .............................................................................................................................. 48

CHAPTER IV: RESULTS ............................................................................................................ 49

Introduction .......................................................................................................................... 49
Description of the Sample ..................................................................................................... 49
Details of the Analysis ........................................................................................................... 49

Research Question 1: What impact does concept map training have upon the complexity of students’ concept maps? .............................................................. 49

Research Question 2: What changes do students exhibit in their concept map characteristics throughout a 16-week science course? ........................................................... 51

Research Question 3a: What impact does concept mapping assessment design have upon students’ concept maps? Specifically, what differences exist between less-directed and more-directed approaches to concept mapping? .......................................................... 52

Research Question 3b: What impact does concept mapping assessment design have upon students’ concept maps? Specifically, what impact does the focus question have on a concept mapping task in terms of the knowledge complexity and content of the resultant concept map? ............................................... 55

Research Question 4: What differences exist between students with high scoring concept maps and students with low scoring concept maps in their structural knowledge, as represented by concept maps? .............................................................. 59

Research Question 5: What differences and similarities can be observed between student concept maps intended to assess structural knowledge and student responses to an open-ended assessment intended to assess content knowledge? .................................................................................................................. 61

Summary .............................................................................................................................. 63
Table of Contents - Continued

B. NOS Quiz Concept Map ........................................................................................................... 87
C. Mitosis (Directed) Concept Map .......................................................................................... 89
D. “What is biotechnology?” Concept Map ............................................................................... 90
E. Views of Nature of Science (VNOS) ................................................................................... 92
F. Human Subjects Institutional Review Board Letter, 17-03-21 ........................................... 96
LIST OF TABLES

1. Assessments implemented per research question…………………………………………..38
2. Pre-instruction VNOS analysis comparing NOS views per tenet…………………………63
3. Post-instruction VNOS analysis comparing NOS views per tenet…………………………64
LIST OF FIGURES

1. Scoring of concept map (#169)……………………………………………………………………..45
2. Student’s first concept map (#76) showing hub/spoke structure………………………………46
3. Effects of concept map training on structural component means ± SD (n=34)……………50
4. Comparison of pre-training and post-training structural shape frequencies (n=34)…………………………………………………………………………………………………51
5. Comparing structural components using the same focus question for post-training “What is science?” Map 2 through 8. Values are means ± SD (n=114)…………………………………………………………………………………52
6. Percent changes between “What is science?” post-training concept maps (n=114)………………………………………………………………………………………………….53
7. Comparison of structural components of a less-directed to a directed map. Values are means ± SD (n=32)…………………..……………………………………54
8. Comparing the integration of a directed mapping task among serial maps. Values are means ± SD (n=44)………………………………………………………………55
9. Comparing two less-directed assessment designs, one with a wide focus question and another with a narrow content focus. Values are means ± SD (n=50)………………………………………………………………………………56
10. Comparing total frequencies and diversity of NOS concepts among two types of less-directed assessment designs (n=50)…………………………………57
11. Comparing structural components using different focus questions. Values are means ± SD (n=33)……………………………………………………………………58
12. Comparing NOS concepts between two less-directed, but different focus question maps (n=33)……………………………………………………………………59
13. Significant differences exist between all structural components between low and high scorers (p < 0.001). Values are means ± SD (n=120)…………………………60
14. Comparing NOS concepts between low-scorers and high-scorers (n=120)…………….60
15. NOS concepts identified in post-training “What is science?” maps (n=114)…………..62
CHAPTER I: INTRODUCTION

Background

Knowledge domains can be organized into structures that show relationships or connections between concepts (Novak & Cañas, 2007). Knowledge about individual concepts develops over time using prior knowledge and relationships with other contexts (Novak, 1993; Novak, 2005; Wehry, Algina, Hunter, & Monroe-Ossi, 2008). Various types of assessments have been used to assess what students know about a particular concept throughout history (Stiggins, 1991; Wilbrink, 1997). Recent theory suggests that there are two main types of assessments, summative and formative. Summative assessments provide information regarding classroom outcomes at an endpoint of the learning environment (Harlen & James, 1997; Stull, Varnum, Ducette, & Schiller, 2011), and therefore summarizes what learning occurred (Harlen, 2005). Formative assessments provide feedback to the teacher or the student and is intended to help learning (Harlen, 2005). Using feedback, students can modify their work or understanding and teachers can modify their teaching (Black & Wiliam, 1998b; Harlen, 2005). Therefore, assessments can gauge students’ knowledge, monitor relationships or connections between concepts, and provide feedback to both teachers and students. While both types of assessments can provide relevant information to a teacher, the overall purpose and timing of the assessment should be carefully considered.

The overall purpose of an assessment should be carefully considered because different assessments serve different purposes of monitoring students’ knowledge. For example, either a formative or a summative assessment, has the capability to inform teachers of misconceptions held by students. However, with a formative assessment, teachers are able to reteach the
concepts to improve student outcomes. In addition, some assessments focus solely on concept acquisition, but do not monitor the relationships or connections students make between concepts (Lazarowitz & Lieb, 2006; Markham, Mintzes, & Jones, 1994; McClure, Sonak, & Suen, 1999; Weurlander, Söderberg, Scheja, Hult, & Wernerson, 2012).

Concept mapping is an assessment tool used to monitor connections or relationships learners make between concepts, as represented in a knowledge structure (Markham et al., 1994). Concept maps have been used to measure learning (McClure et al., 1999; Wallace & Mintzes, 1990), and provide an assessment of new or missing concepts (Burrows & Mooring, 2015; Martin, Mintzes, & Clavijo, 2000). Feedback from concept maps about students’ knowledge structures help teachers and researchers better understand how to promote and monitor cross-cutting relationships or connections that are especially prominent in science (Burrows & Mooring, 2015; Martin et al., 2000).

Statement of the Problem

Assessment has been used for many years to monitor students’ knowledge (Stiggins, 1991; Wilbrink, 1997). There are a variety of assessments that can be used to assess students’ knowledge. However, researchers have found that assessments do not always articulate how students form relationships or connections among the concepts learned in the class (Lazarowitz & Lieb, 2006; Markham et al., 1994; Weurlander et al., 2012). Thus, to capture all the facets of students’ knowledge, multiple assessments are used to better understand the knowledge structures of a student (Burrows & Mooring, 2015). Knowledge structure assessments will vary in their design and function which will impact what aspects the knowledge structure the assessments can capture of the student (Brandstädtter, Harms, & Großschedl, 2012; Ruiz-Primo, Shavelson, Li, & Schultz, 2001).
Researchers agree that student knowledge can be monitored using structural knowledge maps like concept maps, mind maps, conceptual diagrams, and visual metaphors (Eppler, 2006). Using concept maps to monitor students’ knowledge structures help researchers analyze how concepts are related to each other with relationships and connections. Structural knowledge in the form of a concept map can be less-directed, where the student provides the concepts, link phrases, and that will be mapped into their designed structure. This method provides the most insight into what the student perceives are the concepts relating to the focus question being asked (Brandstädter et al., 2012; Yin, Vanides, Ruiz-Primo, Ayala, & Shavelson, 2005). Furthermore, students may be asked to create a map or fill in a map with given terms or linking phrases (propositions). This type of mapping is considered a closed-ended (high directed) concept map (Brandstädter et al., 2012; Yin et al., 2005). Through these multiple connections, researchers can better understand how novice and expert students relate their knowledge (Burrows & Mooring, 2015). A novice map will be less connected among the concepts, whereas the expert map will be well connected among the concept. Regardless of the map reflecting expert or novice knowledge, the overall concept map is never completed (Novak & Cañas, 2007). By evaluating knowledge structures, researchers can begin to evaluate students’ structural knowledge using the content and structure of the map (Markham et al., 1994).

Furthermore, the literature on concept mapping and knowledge structure focuses upon the changes in students’ structural knowledge using an intervention that is marked by a pre-assessment and a post-assessment. While that shows the initial and final changes of the knowledge structure, there is little known about how the knowledge structure changes occur throughout a semester using a the same and a variety of concept maps. Specifically, little known about how concepts in a knowledge structure evolve over time. This study will begin to address
what changes occur in knowledge structure content and complexity between the start and finish of a semester course.

Purpose of the Study

The purpose of the study is to describe how concept mapping can be used as an assessment to monitor what goes on between pre-assessments implemented at the beginning of a course and post-assessments implemented at the end of the course. For this research, concept mapping was used as a serial assessment implemented at various times during several class sessions. The connections (links and linking phrases) students make between concepts suggest possible relationships or connections those students are creating between the concepts (Novak & Cañas, 2007). The overall map structure and quality can be evaluated using various quantitative measurements (Afamasaga-Fuata’I & Reading, 2007; Martin et al., 2000). These measures include counting levels of hierarchy, concepts, links between concepts, and occurrences of merging and branching (Markham et al., 1994; Martin et al., 2000; Novak & Gowin, 1984). In addition, concept maps will be evaluated using the overall structure complexity (Meagher, 2009; Yin et al., 2005). Collecting and analyzing consecutive concept maps will also help address a gap identifying whether students’ overall structural knowledge of science can be mapped to provide insight into their conceptual understandings of science.

Significance of the Study

The research is important to evaluate what changes may occur in a students’ structural knowledge through time. Research has established that pre-assessments and post-assessments are used to monitor what changes occur in students’ knowledge. However, few studies consider what assessments are best used to determine why changes between the pre-post assessments occurs. In addition, when monitoring these knowledge changes between pre-post assessments, limited
research has been found that monitors how students’ structural knowledge changes. Using concept maps to assess what changes in students’ structural knowledge changes as a result of learning has been well documented, but no published research study has had more than four concept maps completed in a semester (Martin et al., 2000; Quinn, Mintzes, & Laws, 2003-04). This research will collect multiple concept map assessments about the same content topic. Additionally, different concept mapping task demands may highlight various reports of knowledge structures. Thus, the research will be useful for understanding how the assessment design can impact the structural knowledge represented by a student.

Theoretical Framework

Studying structural knowledge in the form of concept mapping interweaves the research of Novak with the theoretical work of Ausubel (Novak & Cañas, 2007). David Ausubel, a psychologist, proposed the meaningful learning theory (Novak, 2002). This theory states that an individual learns when new, acquired assimilated knowledge is integrated into pre-existing knowledge (Ausubel, 1963; Novak & Cañas, 2007). This learning theory is also known as Assimilation Theory. Ausubel suggests that learning is either meaningful or rote (Ausubel, 1963). Meaningful learning is exhibited when concepts are integrated within the cognitive [knowledge] structure, whereas, rote learning is when concepts remain isolated within the cognitive [knowledge] structure (Ausubel, 1963). This learning theory uses advanced organizers to monitor the cognitive [knowledge] structure of students’ learning.

In his 12-year, longitudinal study that started in the 1960s, Novak transformed children’s responses from interview data into a structural form called the concept map (Novak, 2005; Novak & Cañas, 2007)). Using the transformed interview data, Novak was able to monitor how students added concepts to their structural knowledge over the 12 years (Novak & Musonda,
1991). Thus, Novak was able to observe progressive differentiation as suggested from Ausubel’s theory, which suggests how subject matter is organized in a discipline. This principle proposes how the content should be organized or how teachers can present materials using a general to specific approach (Ausubel, 1963, 1970). Thus, a teacher provides an overview to a topic then progressively reveals more specific content. This overall progression from more inclusive to less inclusive topics mimics the principle of progressive differentiation, but has not always been taught in classrooms (Ausubel, 1963).

The second principle of Ausubel’s meaningful learning theory is integrative reconciliation. Ausubel uses a textbook to illustrate how ideas can be segmented. Most textbooks have topics of knowledge segregated into chapters and portray a limited view of integrating particular ideas (Ausubel, 1963). This compartmentalization, or segregation of topics, will lead to similar ideas being represented in a variety of ways and given terms with few relationships. Per Ausubel, more effort should occur reconciling what relationships or connections exist among topics and concepts. Thus, meaningful learning results when concepts can be related or not related and reconciliation must occur to explain the relationship (Novak, 2002). Thus, the meaning of concepts is represented by the linking words (propositions) that integrate the concept into the structural knowledge framework.
CHAPTER II: REVIEW OF THE LITERATURE

Structural Knowledge

When defining “knowledge” there are epistemological factors that should be considered in respective content domains. In general, there are two main types of knowledge that many researchers agree upon: declarative knowledge and procedural knowledge (Jonassen, Beissner, & Yacci, 1993). “Declarative knowledge (called conceptual knowledge) is knowledge where we know that about something, whereas procedural knowledge is where we know how something works” (Novak, 2002, p. 553). Specifically, declarative knowledge constitutes facts, definitions of concepts, and descriptions, whereas procedural knowledge represents the rules or sequences of the domain specific content (Shavelson et al., 2005). In distinguishing differences beyond this, there are five descriptive characteristics that should be considered when evaluating types of knowledge: level, structure, automation, modality, and generality (de Jong & Ferguson-Hessler, 1996, p. 111, Table 1).

Jonassen et al., (1993) have suggested that “structural knowledge” is the transformation of declarative knowledge into procedural knowledge. Thus, structural knowledge is the “conceptual basis for why; it describes how the declarative knowledge is interconnected” (Jonassen et al., 1993, p. 2). Often, researchers have described structural knowledge as a cognitive structure, “the pattern of relationships among concepts in memory” (Preece, 1976). Thus, in this comprehensive and critical review, structural knowledge will reflect how someone organizes the concepts in their own memory and how those concepts may change the overall knowledge structure. When learning occurs, more concepts are added to the overall knowledge
or cognitive structure. The resultant structure is a function of meaning as it is associated with memory.

Monitoring knowledge structure may be of interest for instructors because they would be interested in what their students are learning from the content being taught. Instructors may be interested in how their students are assimilating the material to their current understanding or knowledge structure. Thus, the need to have continuous opportunities to monitor knowledge is necessary. One particular way that instructors can assess learning is by asking students to visualize their own knowledge structures.

Knowledge Structure Construction

Assessing knowledge construction using concept maps has also been used as a learning tool and as an evaluation tool in research (Novak & Cañas, 2007). Knowledge defined as being “stored in our brain [and] consists of networks of concepts and propositions” (Novak, 2002, p. 551). Just as the many ways to learn is diverse and unique, there are many ways to capture knowledge and the strategies are not fool-proof (Novak, 2002). Just as learning has no set of fool-proof methods, eliciting that knowledge can also be difficult. However, insight of the changes that occur as a result of instruction may be useful in exploring how students are learning and what accommodations are being made when they learn new knowledge (Shavelson et al., 2005). Researchers have monitored such restructurings of knowledge by using concept maps (Martin et al., 2000; Meagher, 2009; Quinn et al., 2003-04).

A theme that came out of the knowledge structure literature is the opportunities to observe knowledge structures while they are created by an individual. Eppler (2006) provided comparisons of how various tools can be useful for demonstrating and sharing knowledge construction. Eppler (2006) compared concept maps, mind maps, conceptual diagrams, and
visual metaphors in terms of definition, main function or benefit, typical application context, application guidelines, employed graphic elements, reading direction, core design rules or guidelines, macro structure adaptability, level of difficulty, extensibility, memorability, understandability by others, and typical software package supporting the visualization format (Eppler, 2006, p. 203-204, Table 1). The author analyzed the literature and could make comparisons of four advantages and four disadvantages descriptors among the four knowledge structure tools (concept map, mind maps, conceptual diagrams, and metaphors) being addressed in the paper (p. 206, Table 2). This table provided a summarized view of thirteen research studies of their main advantages or disadvantages as it related of how they studied one of the four knowledge structure tools. Researchers and teachers may find this table be useful in their own research in considering what tool would be most appropriate for their research.

Researchers may only a single concept map to represent student knowledge (Burrows & Mooring, 2015; Markham et al., 1994), but both papers utilized interviews to better understand the meaning of the students’ knowledge structure. Whereas, Ketpichainarong, Panijpan, and Ruenwongsa (2010) used concept maps as a pre- and post assessment to better understand change in student knowledge structure as a result of an intervention of an inquiry biotechnology lab. While concept mapping was not the only tool to monitor the intervention (conceptual understanding test, laboratory group report, Constructivist Learning Environment Survey (CLES) questionnaire, and student interviews), the authors did use concept maps to monitor changes to the overall student understanding of biotechnology knowledge. Other researchers used concept mapping as a tool to monitor changes of their cognitive structure throughout the entire semester (Martin et al., 2000; Meagher, 2009; Quinn et al., 2003-04). Martin et al., (2000) collected four concept maps (week 1, 6, 12 and 16) and completed as series of interviews where
students were asked about their concept maps. Meagher (2009) collected three concept maps over a semester. Quinn et al., (2003-04) collected three maps during the 15-week semester and another concept map with a sub-set of participants 6 months after the class ended. Each of these researchers used a different methodology to monitor structural knowledge in their classrooms.

Structural Knowledge Map Contributions

Researchers have established that knowledge can be represented in written artifacts (Eppler, 2006). Their theoretical background for understanding structural knowledge is derived through constructivist theory, which suggests that people construct their own meaning through personal experiences and reflections of such experiences. Constructivist theorists suggest that the analysis of knowledge structure can be informative of declarative knowledge. Novak (2002) provides a definition of knowledge that aligns with learning and capturing knowledge. Thus, major trends of the structural knowledge research are supported by mapping structural knowledge. Identifying the connections can be useful in how a participant was relating concepts to each another.

Connections among various concepts will show a teacher or researcher understand how the concept can be successfully or unsuccessfully integrated in the knowledge structure (Martin et al., 2000). When the restructuring of knowledge can be monitored, it is useful in identifying key concepts and misconceptions, and lack of expected concepts. This information provides teachers and researchers information about curricular changes and learning progressions. Furthermore, how they approach the identification of concept arrangement and connections may need more research.
Structural Knowledge Map Limitations

Researchers agree that structural knowledge can be mapped by teachers, students and researchers using a variety of written structures to represent knowledge (Safayeni, Derbentseva, & Cañas, 2005; Shavelson et al., 2005). However, different mapping techniques provide different aspects of knowledge being mapped (Shavelson et al., 2005). The variety of the tasks may make it difficult to synthesize the results and can also lead to misinterpretation of the maps using their qualitative (overall structure) and quantitative methodologies (Ruiz-Primo et al., 2001). A variety of maps may occur when giving a task where the students are asked to create the entire map without any guidance. However, this type of map design may provide the most information of how a student understands the concepts of interest (Brandstädter et al., 2012; Yin et al., 2005). Furthermore, different constraints of a mapping task can limit what information can be interpreted from the map (Ruiz-Primo et al., 2001). A mapping activity that asks the student to fill in concepts from a word bank will suggest a different student understanding than having a student create their own map from their own concepts (Brandstädter et al., 2012; Yin et al., 2005). Just as giving students linking words to connect the concepts is a more directed approach than asking students to create and place their own linking words among concepts. A variety of designs can be used when designing a mapping activity.

While having students, teachers and researchers document their own structural knowledge, there are research limitations for younger students who cannot articulate or have limited ability to write. Researchers have interviewed elementary students regarding their understanding and constructed their knowledge from interview tapes (Wehry, Algina, Hunter, & Monroe-Ossi, 2008). Their methodology was similar to how Novak constructed his 12-year longitudinal study (Novak & Musonda, 1991). Wehry et al., (2008) interviewed preschool age
children of their knowledge about plants and constructed concept maps from their interview transcripts. However, there has been limited research of how elementary and younger students’ knowledge is structured because of language and written demands per the task.

Finally, eliciting and revisiting student knowledge structures can be time-consuming in a classroom. Many researchers have agreed that obtaining this information can be useful for the student and the teacher (Burrows & Mooring, 2015; Eppler, 2006; Ketpichainarong et al., 2010; Markham et al., 1994; Martin et al., 2000; Meagher, 2009; Wallace & Mintzes, 1990; Wehry et al., 2008). Limitations of time in the classroom are a reality of this methodology being implemented in a variety of classroom and research settings.

Summary

As previously stated, concept mapping can be useful to better understand or iterate known relationships or connections between concepts in a knowledge structure. Thus, when students are asked multiple times throughout the course to create a concept map, the changes that can be observed may provide insight into their initial or advanced understandings of the content. Furthermore, multiple concept maps can be a unique way to gain insight into a few key concepts.

Finally, concept mapping can be helpful in making connections within the knowledge structure of an expert or novice. The literature supports the idea that expert maps are more complex and more inter-related among concepts (Burrows & Mooring, 2015). Often, the novice maps are linear paths or concepts with few connections like a spoke map. Comparing an expert map to a novice map may be helpful to identify the missing connections and concepts that exist between a group of experts and a group of novices within a particular field or classroom.
Assessment in a Classroom

Historical Context of Assessment

Reflecting on how learning assessments have been implemented historically can help researchers and teachers better understand how assessments are used today. Through the 5th and the 15th centuries, also known as the medieval times, assessments focused on practicing and recitating spiritual or sacred texts that were then to be passed down familial lineages (Wilbrink, 1997). This particular method of recitation was a method to learn difficult texts. Many of the texts were written in other languages which also had to be learned. Instructors asked students questions and students were assessed based on their ability to recite the correct, memorized responses. To memorize and recite a text was to know it by heart (Wilbrink, 1997, p. 32). This assessment approach evolved into the common lecture style of teaching that universities and colleges commonly use today.

Another example of a historical assessment was disputation which also occurred during the medieval period. The disputation was “a theorem or problem posed by the master” (Wilbrink, 1997, p. 36). Students were required to orally defend the master’s (author’s) work and their responses “could be opposed by the masters and the [other] students” (p. 36). This examination format was important for the new development of various styles of question and answer techniques. There was an overall motivation to do well by the student being examined and to represent the work of the master well. Recitation and disputation are similar to how current policy holders and teachers will publicly have to report what their students are learning (Wilbrink, 1997, p. 37). Disputation was an earlier form of teaching and learning which then evolved into the period of catechism. Catechism is learning through questioning and answering of a religious text. While the catechism practice was used as a form of memorizing historical and
religious texts the social and economic factors began to change. Industrialization created an influx of diverse learners that caused the schools to accommodate this number of learners and diverse assessment methods (Stiggins, 1991). The increase of diverse learners pushed instructors to incorporate less recitation and more paper and pencil testing into their assessments, thus assessments were starting to vary based on the needs of the schools (Stiggins, 1991, p. 265).

In addition to the influx of learners and more methods of assessment, teachers were facing the use and misuse of punishment for the students’ undisciplined behavior (Wilbrink, 1997). The amount of harsh punishment was reviewed by humanists who suggested a positive behavior and prize marking system in the early 19th century. Students were ranked by their merit and achievements of that school year. Many schools in England used this ranking and merit system to honor their students at the university level. The scores that students received seems to have been transformed from a merit based system into a marking system. However, various countries have different marking systems. This particular history is important to consider and evaluate when the Western world identifies and relates the expectations of assessment to teachers, students and parents.

Competency testing has been used to support how teachers assessed as they “saw fit” (Brookhart, 2013) This movement to paper and pencil assessments led to standardized testing in the 1930s for select high school students. These standardized tests became the foundation for college admittance testing in the 1940s (Stiggins, 1991). This new college standardized testing was aligned with the objectives from the 1930s (Brookhart, 2013). From the 1950s to 1970s, more emphasis was placed upon learning objectives and their relationship with influencing minimum competency (Brookhart, 2013). Also, “it became clear that minimum competency testing lowered expectations to meeting minimum requirements” (Brookhart, 2013, p. 70). In
addition, federal funding for educational initiatives was impacted by the idea of standardized tests. Specifically, in 1965 the Elementary and Secondary Education Act (ESEA) provided a set of guidelines to be used to determine how five billion dollars were spent in funding education (Anderson, 2012, p. 106). All monies given came with accountability tasks, including measurable changes in assessment scores. An act like ESEA that funds education and requires assessment accountability lays the foundation for more policies that entwine education funding and assessment accountability to the discourse of the accountability systems that schools use today.

During the 1960s, psychology and the philosophy of science was also being influenced by the constructivist movement in education (Klassen, 2006). After this empirical and behaviorally dominated learning paradigm, cognitive psychology and constructivism blossomed through the work of Kuhn and Piaget (Klassen, 2006). Understanding the theoretical views of knowledge clarified what we understood as the theoretical origins of assessment. Furthermore, the distinguishable differences between the knowledge structure of a scientist and a science student can influence future work in science education (Klassen, 2006, p. 828).

In 1983, a national publication ‘A nation at risk’ was available by the National Commission on Excellence in Education (NCEE, 1983) which highlighted how US students were underperforming in science and technology (Brookhart, 2013). This led to an educational reform in the 1980s and 1990s that encouraged states to develop educational standards. However, the states varied in the standards and policies they had adopted. To account for these differences the federal government legislated the No Child Left Behind Act of 2001 (NCLB, 2002) which “…reported proficiency level on standards, as measured by standardized tests to the federal government” (Brookhart 2013, p. 71). Recently, many states have adopted the Next
Generation Science Standards (NGSS) which were collaboratively created for implementation at the K-12 level (NGSS Lead States, 2013). It is anticipated that these changes at the K-12 level in science education will feed into college level programs and teacher education programs, thereby impacting all levels of education.

Types of Assessment

The National Research Council (NRC, 1999) and Scriven (1967) have delineated two types of assessment: summative and formative assessment, based upon their purpose or intended use. Formative assessment is intended to help learning and summative assessment is intended to summarize learning (Harlen, 2005, p. 208). Researchers have also used the phrase, “assessment of learning” to describe summative assessment and “assessment for learning” to describe formative assessment (Harlen, 2005). Bell and Cowie (2001) reinforce the idea that the purpose of the assessment will help define the type of assessment (p. 537). A current definition of summative assessment has been clarified as “summing up or summarizing the achievement” (Sadler, 1989, p. 120). Formative assessment uses feedback to drive learning practices for both the teacher and the student, as well as to drive knowledge development for the student (Harlen, 2005).

Summative assessments have been utilized at all levels of education to summarize student learning. More often, a summative assessment is taken by the student and is an end point to a sequence of learning events (Stull et al., 2011). Assessment near the end of the learning unit has little impact on the continuation of learning for the student. Harlen and James (1997) suggest that summative assessment has particular criteria that will relate students’ assessments against “other interested parties” (p. 370). Instructors assess students using specific criteria or objectives that the student must meet. Individual student results can be combined since they are “based on the
same criteria” (Harlen & James, 1997, p. 373). Often the standard unit of comparison for some teachers are multiple selected responses (Klassen, 2006). This type of assessment yields a quantitative product of what learning has occurred for each student.

The NRC (1999) identified summative assessment as having multiple purposes that use an assessment of the standards to compare student scores to each other. Summative assessments have also been used to provide reports to parents, schools and national assessors (Harlen & James, 1997). Employers and higher education institutions can also find students’ summative assessments useful (Harlen, 2005, p. 208). Summative assessments are used in higher education to report in their accreditation of academic programs, Also, assessments can be used by providing a summary of students work for students’ future internships and employers.

Summative assessments are further differentiated as being ‘internal’ or ‘external’ (Harlen, 2005). Internal summative assessments utilize the assessments to inform the student, parents, or are used for regular (classroom) grades. External summative assessments are used for certifications and school performances. The purposes of external summative assessment extend beyond informing the teacher or the student of progress. Teachers may have to use the summative assessments to support school or district data collection in regards to teacher performance and overall school scores (Harlen, 2005). Additionally, summative assessment has been used as a way for policy makers to use assessment data for accountability in educational reforms (Harlen, 2005; NRC, 1999).

Scriven (1967) initially distinguished between summative assessment and formative assessment (Wiliam, 2006, p. 284). However, Sadler (1989) first provided the theory of formative assessment. Meanwhile, Black and Wiliam (1996a) are also recognized as major contributors to formative assessment theory. However, some researchers have pointed out that
Black and Wiliam’s work lacked empirical research (Kingston & Nash, 2011). Regardless, other researchers have used Black and Wiliam as theoretical framework for their own formative assessment research (Bell & Cowie, 2001; Harlen & James, 1997; Haug & Ødegaard, 2015; Lazarowitz & Lieb, 2006; Sadler, 1989; Weurlander et al., 2012).

Black and Wiliam (1998a) defined formative assessment as, “all those activities undertaken by teachers and/or by their students, which provide information to be used as feedback to modify teaching and learning activities in which they are engaged” (p. 7). Feedback and modification of teaching and learning are the key aspects of formative assessment. Both the student and the teacher may provide feedback or learn something at any one time. However, when no further feedback is given and an assessment is given, the assessment would be considered summative.

Formative assessment is a task that provides opportunity to elicit feedback from the students to the teacher or from the teacher to the students. Teachers and students work closely together “to modify their work in order to make it more effective” (Black, 1993, p. 49). Harlen and James (1997) suggest that the teacher and student will provide feedback of each of their understandings to “determine the way forward” (p. 369). Formative assessment is a cycle of information that will be informative of what learning has or has not occurred through a process of feedback.

Formative assessment informs the student and the instructor of the conceptions the student has in their knowledge domain. Students completing a formative assessment will provide instructors feedback about the student’s own knowledge. This student feedback from their assessment may help determine what next action the instructor should take to address future objectives and what additional follow-up is required with the student. Thus, students use a
formative assessment to iterate to the instructor what prior knowledge they have. While prior knowledge is helpful to the instructor to impact the future course objectives, formative assessment can also provide self-awareness to the student. The student’s awareness highlights what content and connections were understood and what content aspects should be reviewed and focused upon for clearer understanding (Weurlander et al., 2012). Thus, formative assessment can be a tool or method that an instructor extends to their students to learn the content requirements of the course.

Teachers Impacting Assessment

To create effective formative assessment, Haug and Ødegaard (2015) suggest that teachers require a high level of pedagogical content knowledge (PCK; Abell & Siegal, 2011, p. 206). If a teacher has a limited PCK, they are “less likely to know what questions to ask students to elicit their ideas of which conceptual difficulties to anticipate” (Haug & Ødegaard, 2015, p. 633). Teachers who have a high level of PCK are able to understand and assess relevant content knowledge, idealize the best timing of the assessment, be able to interpret the results or feedback, and redirect or reinstruct, if necessary.

Integrating a variety of assessments options for students can be very time consuming (Craddock & Mathias, 2009). Teachers need to consider the “what and where” to implement formative assessments in their classroom (Ayala, Shavelson, Ruiz-Primo, Brandon, Yin, Furtak, Young, & Tomita, 2008). Researchers have developed a guiding framework using six categories to map out assessments using the Assessment Design Decisions Framework (ADDF) (Bearman, Dawson, Boud, Bennett, Hall, & Molloy, 2016). The ADDF was created to acknowledge six factors that influenced assessment design and would interact to create a more learner centered approach to assessment (p. 551). This was a published framework to encourage educators to
consider all factors when designing and implementing assessment. Additionally, educators have utilized and published feedback loops (Furtak, Glasser, & Wolfe, 2016) and pre-designed formative assessments for a science classroom (Keeley, 2015). However, more research is required to better understand the limitation of how teachers utilize their PCK with their assessment literacy in creating assessments for classroom learning.

Sometimes, teachers have the PCK to create relevant formative assessments, but are limited in their ability to implement those assessments effectively. Some of these limitations may come as a direct result of the pressures of high stakes testing. High stakes summative assessments have been the priority in many classrooms rather than open-ended, student centered learning (Marsh, 2007). However, Haug and Ødegaard (2006) found that “classroom talk was a preferred method for observing students’ use and understanding of new science words. However, a written test was considered best, and, for some, the only way to collect information that provided valid information on student learning” (p. 642). Furthermore, Black, McCormick, James, and Pedder, (2006) suggest that more schools may integrate formative assessment strategies if the strategies were optimized for a more generalized application. Black and Wiliam (1998b) recognize that the implementation of formative assessment can be slow for a teacher, but long-term engagement of implementation can improve teaching and learning within the classroom.

Assessment Impacting Students

A student can be impacted by the implementation of formative assessment by self-assessing their own learning against the objectives of the course. This awareness of their own learning by completing an assessment or providing feedback causes the student to reflect upon how other applications or contexts may be used to apply the material. This awareness of their
own thinking is often referred to as metacognition. Nicol and Macfarlane-Dick (2006) suggest that good feedback practice relies upon seven principles:

“(1) helps clarify what good performance is; (2) facilitates the development of self-assessment (reflection in learning); (3) delivers high quality information to students about their learning; (4) encourages teacher and peer dialogue; (5) encourages positive motivational beliefs and self-esteem; (6) provides opportunities to close the gap between current and desired performance; (7) provides information to teachers that can be used to help shape teaching” (p. 205).

A student that is provided an opportunity to reflect may extend the opportunity to further impact their peers’ learning. Influences on peer learning may occur by providing relevant feedback to the instructor about their own learning where the instructor may use the feedback to redirect or reinstruct the student, a group of students, or the entire class. This feedback requires active participation from the student(s) and is part of the culture of a classroom (Craddock & Mathias, 2009). Thus, the feedback can cause the student to provide feedback to themselves or the teacher of where they are in the conceptual understanding of a particular topic. Successful formative assessment allows students to connect concepts to contexts while affirming their understanding of the topic.

Influences of Assessment

Researchers would agree that formative assessment can inform the teacher of how and what instruction (or lack thereof) can impact the student’s knowledge structure. This overall impact of teacher instruction has been well-studied and researchers suggest a variety of ways to teach for the maximum impact of student learning (Bell & Cowie, 2001; Black, 1993; Harlen & James, 1997; Wiliam, 2006). However, teacher PCK and classroom management can strengthen
how formative assessment is utilized in a classroom. Thus, when reviewing the impact of formative assessment in a classroom, researchers and teachers should take note of this aspect. One purpose of formative assessment is to expose the learning that occurred and can impact current teaching and future teaching. However, limited research has been done on what aspects of content or assessments have been updated because of specific formative assessment.

Researchers would also agree that when a student participates in formative assessment, they become more aware of their own learning or what they may have not yet learned (Bell & Cowie, 2001). Thus, formative assessment can highlight to the researcher, teacher, and student how and what the students are learning. This awareness may extend their own learning to influence their peers’ learning. Unfortunately, this sharing experience may be the result of how the classroom is managed by the teacher, rather than from the impact of formative assessment itself.

Finally, researchers have not been able to articulate what formative assessments may be influential in changing summative assessments. This may be the result of the literature that was reviewed for this literature review as it was for college level classes. In a college class, similarly to other classes, limited time is allowed for formative assessments and measuring its impact to the final assessments. Thus, more research is necessary to correlate any types of formative assessment research to summative assessment research. More research would be necessary to observe what impacts one assessment versus the other assessment.

Assessment Limitations

Researchers and teachers utilize assessments to evaluate student knowledge. However, their assessment will depend upon the PCK of the researcher or the instructor implementing the assessment (Haug & Ødegaard, 2015). Limited information is provided in research articles about
the content knowledge of the researcher and teachers implementing assessments. More information should be provided in the literature about teacher and researcher PCK to better understand if any weaknesses may exist in their methodology in using assessments. Thus, more research is necessary to understand how teacher PCK can influence assessment literacy, implementation of assessment, and linking implemented assessments to learning.

Researchers and teachers have delineated among assessments as summative and formative (Harlen, 2005; Scriven, 1967). Summative assessment provides limited information to modify future teaching or learning because summative assessment may occur near the end of the unit. However, the influence or weight of the summative assessment relating to outside classroom goals of reporting grades, provides limited opportunities to explore formative assessment to monitor student understanding (Harlen, 2005).

Researchers and teachers use formative assessment as a learning tool and not as a graded assessment. There are a variety of assessments that researchers and teachers have used formatively (Keeley, 2015). In addition to mapping the learning progressions of science, formative assessments can benefit students by providing feedback about current understanding of the topic and be a tool for learning content and motivating students to study the material. The type of feedback that can be useful for the research or teacher can also be limited by the variety of formative assessments implemented. Feedback requires active participation of the instructor and/or the student to complete the feedback loop (Craddock & Mathias, 2009). Limited studies have monitored the complete feedback loop well in their classroom (Lazarowitz & Lieb, 2006; Weurlander et al., 2012). Furthermore, limited research has shown what happens when students use the tools and still cannot grasp the content. Or, when students understand the content but are not motivated. More research is necessary to distinguish these difficult parameters as they relate
to formative assessment. Future assessments would require students to be transparent of their use of formative assessment, lack of motivation, and content knowledge.

Conclusions and Research Gaps

Formative assessment in a science classroom can benefit both the instructor and the students. Models that suggest the importance of pre-test formative assessments can improve instructor responses to students’ lack of knowledge or misconceptions of science concepts. Just as individual completes written formative assessments for insight into their own knowledge, group formative assessments can also help guide the student in being aware of content knowledge (Weurlander et al., 2012). Furthermore, these assessments can help expose what concepts students should focus on and study. The results of an assessment may encourage and motivate the student to apply their knowledge to other concepts and continue to learn more about the topic being taught and learned. However, more research is necessary in determining how the methodology of the type and timing of assessments may be integrated within science curriculum.

Summary

While exploring implementations of formative assessment in a science classroom, research is clear that measuring students’ relationships or connections among concepts they are learning is an important aspect of assessment (Lazarowitz & Lieb, 2006; Weurlander et al., 2012). The tools being used in the classroom should reflect measurement tools that can better understand what connections are being made by students. Thus, assessment tools help provide feedback to the researcher or teacher of how to approach the next steps of meaningful learning. Therefore, researchers and teachers alike should be careful implementing assessments that can measure their cognitive level but can also show how students relate concepts together.
Concept Mapping Knowledge Structures as Assessment

Concept mapping is a tool that can be implemented in research methodologies and classrooms to review or evaluate what connections are being made among concepts. Sometimes, the concept map will answer a focus question that will help associate the connections between concepts (Novak & Cañas, 2007). A concept map is a drawing of connected concepts. The concepts are connected using arrows or links that connect the circled concepts. On the arrows or links, the map creator will include short phrases often called linking words or phrases. The linking words or phrases that connect the concepts will create a knowledge structure of a content area. The concept mapping task may have various requirements that allows the assessment to be directed or less-directed (Ruiz-Primo et al., 2001). Similar terms have been used to define mapping tasks, less-directed may be considered a “free-form” and a high directed (or more directed) map would be considered a “fixed form” (Novak, 1998). For this research, the concept mapping tasks will be discussed as less-directed and directed as that suggests more of a continuum as shown by Ruiz-Primo et al., (2001). Students may be asked to only provide the concepts or linking phrases of an already structured concept map, which is considered a directed (or more directed) concept map (Ruiz-Primo et al., 2001). Other task variations may ask students to create a map from the given concepts or linking words/phrases. These methods provide more direction and allow for limited interpretation of a map creator’s structural knowledge (Brandstädter et al., 2012; Yin et al., 2005). Furthermore, the task may have less direction, where the map creator should conceive most or all of the map components (Yin et al., 2005). The less-directed approach is having students create a concept map with their own map components. This approach gives the least direction and more freedom to the map creator to represent their structural knowledge of a content area (Brandstädter et al., 2012; Yin et al., 2005).
History of Concept Mapping

Concept mapping was initially used by researchers to monitor students’ science knowledge over time. Novak and his team interviewed students in a 12-year longitudinal study to better understand how a student’s science knowledge changed over time. They used the interview data to draw and connect students mentioned ideas from an interview (Novak, 2005; Novak & Mosunda, 1991). Students’ ideas were transcribed from interview data and were “mapped” by researchers using the students mentioned concepts, linking words/phrases, and links. Through these links, Novak (2005) and his team of researchers could better understand what connections were made between concepts these students were verbalizing during their interview and how they were connecting these ideas. Since Novak’s longitudinal study, educators and researchers from K-16 have utilized concept mapping in their classroom for a variety of reasons including retrieving feedback from students, helping student make connections between concepts, and as assessments (Afamasaga-Fuata’I, 2007; Austin & Shore, 1993; Barney, Mintzes, & Yen, 2005; Beyerbach, 1988; Freeman & Jessup, 2004; Hay, 2007; McClure et al., 1999; Rice, Ryan, & Samson, 1998; Sellmann, Liefländer, & Bogner, 2015; Vanides, Yin, Tomita, & Ruiz-Primo, 2005; Wallace & Mintzes, 1990; Wallace, Mintzes, & Markham, 1992; Yin & Shavelson, 2008; Yin et al., 2005). Thus, concept mapping has been used as a tool to elicit information about structural knowledge (Anohina-Naumeca, 2015; Novak, 2002).

Prior sections highlight how knowledge can be arranged into a structure and the usefulness and strategies of assessment, specifically, summative assessment and formative assessment, in a classroom. This section will highlight key characteristics unionizing the theory and methodology of comparing the knowledge structures of the concept mapping tool to both summative assessment and formative assessment. The concept mapping tool will be compared in
its methodology and usefulness by merging with both formative and summative assessment strategies in a science classroom. Explaining the overall use and how to create concept maps are important for both summative and formative assessment. Concept mapping has been used as a tool to elicit information about structural knowledge (Anohina-Naumecia, 2015; Novak, 2002).

Concept Mapping as Summative Assessment

A summative concept map is used when no more learning that will require feedback will be occurring, thus the concept map is the last task of a unit or class. A researcher using summative assessment through concept mapping would review how the concept maps are scored qualitatively and quantitatively. There are a variety of ways that concept maps are scored (Ruiz-Primo & Shavelson, 1996), but the Novakian system is still used to evaluate concept maps (Novak, 2002). The Novak system scores each of the components of the map (hierarchy, number of concepts, links, linking phrases) for a mapping score for each student. Concept maps can be also scored qualitatively by their overall structure or linkages that occur within the map. There are distinct patterns that researchers look for, they are: linear, tree, circle, hub and spoke, network/wheel (Meagher, 2009; Yin et al., 2005). However, when tools are used within a classroom and if they are being used as assessments, the rubrics for the score and qualitative aspects should be clearly explained and utilized prior to the summative assessment.

A concept map implemented as a summative assessment can be utilized in a variety of ways. Concept mapping can be used as part of an assessment to highlight connections of a variety of contexts or key words. However, if the concept map is being used for a score, instructors should teach the students how to properly use this method of assessment. If instructors do not teach how to concept map, the assessment may not reflect the most valid externalization of a knowledge structure for a student. The reason for this is because they are
limited by the constraints of how to approach and complete the task. One way that instructors have utilized instruction is by providing an opportunity to practice prior to a pre-assessment for the unit or class (Martin et al., 2000; Quinn et al., 2003-04). Also, this pre-concept mapping summative assessment may be useful in comparing to the final or post-concept mapping summative assessment to determine how the teaching, or intervention, influenced knowledge structures. While the concept map summative assessment can be useful for providing a grade for the students’ externalizations of a knowledge structure, more researchers have used concept mapping tools to identify how key concepts from the curriculum are integrated in student knowledge structures in a concept mapping formative assessment (Hay & Kinchin, 2008; Reiska, Soika, Möllits, Rannikmäe, & Soobard, 2015).

**Concept Mapping as Formative Assessment**

Concept mapping may also be implemented as a formative assessment in a science classroom. Teachers use formative assessment in various ways to monitor the quality of student learning and their understanding of scientific literacy (respectively, Hay & Kinchin, 2008; Reiska et al., 2015). Concept mapping is a tool that can reflect knowledge structure which may be useful to gather feedback from the student or from the teacher. Implementing formative assessment as concept mapping in a science classroom may be valuable to understand what applications and connections are being made within science content. Integrating concept mapping may help a student reflect upon their own knowledge connections and interrelationships in their overall understanding of science. Well-connected knowledge structures in the form of a concept map can provide feedback to the teacher of what and how the students externalize their understanding of science content. Additionally, this tool can also be valuable in recognizing any misconceptions and errors in the knowledge structure. However, a teacher’s PCK should have
foundational science and assessment content to be able to implement concept mapping as a formative assessment tool.

Researchers have identified two overall purposes to guide the use of concept mapping as formative assessment. There is a metacognitive purpose that encourages students to reflect upon their learning and suggest improvement to their cognitive skills (Bramwell-Lalor & Rainford, 2014). Anohina-Naumeca (2015) have also utilized concept mapping as an action research approach by having the instructor/author reflect on how to integrate and use concept mapping as formative assessment.

Then, there is the purpose of using concept mapping as formative assessment for the teacher to monitor how the student is connecting various classroom concepts (Reiska et al., 2015). The opportunity to monitor students can be completed over time and can be helpful to suggest what gaps exist in student learning and overall classroom learning (Hay & Kinchin, 2008). A teacher using formative assessment will record and monitor the represented knowledge gaps, misconceptions or altogether wrong content. Using the concept mapping as formative assessment, the teacher can reiterate or address those concerns as they appear in each map or overall.

Extrapolating the knowledge for the concept map structure can be difficult and slow for a student, but over time, the concept mapping process can become easier for the student as they become more comfortable with the task (Cañas, Reiska, & Novak, 2016). However, the experience of creating “good” concept maps can still be difficult (Cañas et al., 2016). The differences in better concept maps reside in the differences in concept maps between novice and expert maps (Burrows & Mooring, 2015; Cañas et al., 2016; Markham et al., 1994).
Major Research Concept Mapping Themes

There were two main themes found within the concept mapping literature. First, concept mapping has been used with other assessments to provide a more thorough understanding about the development of cognitive structure (Burrows & Mooring, 2015; Ketpichainarong et al., 2010; Markham et al., 1994). There are benefits to the teacher and the student if this can be done correctly. This composite of assessments can also have some weaknesses associated with the implementation and analysis. For instance, the absence of timely feedback to the student and the ability to link the various interpretation of a variety of assessments to a learning outcome can be done incorrectly. In addition, the literature of concept mapping has focused on how concept mapping can be used to better understand the differences between expert and novice concept development (Burrows & Mooring, 2015; Markham et al., 1994). Assessing expert-to-novice trends are helpful in developing curriculum maps and finding where the gaps may exist in teacher PCK, curriculum, or standards.

Concept Mapping with other Assessments

There were three articles that were relevant for the contribution in this paper, specifically, in the discussion of using concept mapping in conjunction with other assessments. Burrows and Mooring (2015) had students create a concept map from a given list of terms. The students completed this concept mapping experience during a think aloud interview. Whereas, Ketpichainarong et al., (2010) had students complete a pre-test conceptual understanding test, the Constructivist Learning Environment Survey (CLES) questionnaire and a concept mapping exercise prior to an inquiry-based lab unit. Upon completion of the lab unit, the students completed all three tasks again as a post-test. Additional artifacts were collected from students that included a laboratory group report and student interviews. Markham et al., (1994) utilized
concept mapping with a card sorting task. These articles provide information on the how researchers have used concept mapping to supplement other types of assessment. A variety of assessments can benefit both the teacher and the students by providing the student a variety of opportunities to describe their knowledge. Multiple types of assessments can benefit the teacher by providing a robust picture of students’ knowledge and any misconceptions.

Using multiple assessments can provide awareness of student learning. Students’ acquisition of knowledge will vary on how they acquired the knowledge, when they acquired the knowledge, and what knowledge students acquired. By implementing multiple assessments, researchers and teachers may be provided insight into a student’s structural knowledge change over time (Burrows & Mooring, 2015). Furthermore, the teacher may be provided insight into overall classroom misconceptions of content.

Multiple assessments from several students in a classroom or research setting can provide data necessary to highlight effective or ineffective teaching pedagogies. This measured efficiency aligned to curriculum standards may be useful in correcting and highlighting bad or good teaching, respectively. Furthermore, the mapping of several assessments over time can provide needed information regarding program outcomes or sequences in any science education program (Quinn et al., 2003-04). More work is necessary to better understand how multiple assessments can timely and efficiently align with learning outcomes.

Concept mapping has been used by researchers and teachers in the classroom and for research to better understand how students structure their knowledge. The knowledge may be assessed by the overall structure or the content and links provided. Often, researchers will use concept mapping with other assessments (Markham et al., 1994). However, there has been limited research provided in regards to how the results of the concept mapping directly link to
the results of other assessments. Specifically, researchers and teachers will use various forms and types of assessments, but limited information is provided in regards to how concept mapping can be related to other assessments. However, concept mapping can provide insights into how knowledge is structured and linked. Other assessments (multiple choice tests, likert scales, or open-ended surveys) may just provide information regarding knowledge students may or may not understand (Weurlander et al., 2012). More research is necessary to align how various assessments can be linked to various aspects of concept mapping methodology. More research is necessary to bridge how a teacher or researcher can understand the impact of understanding students’ structural knowledge in various forms of assessments.

Expert versus Novice Aspects of Concept Mapping

Expert versus novice aspects in classroom assessment yield differences of knowledge structure and its representation in concept mapping (Novak, 2002). When someone is an expert of the field, they will have a more complex knowledge structure from their deep knowledge (Shavelson et al., 2005). Abd-El-Khalick and Lederman (2000) published a concept map suggesting that their map as experts in the domain of nature of science would be more complex. This map would be different to a novice map where a novice in a topic area will represent “limited knowledge structures with few connections and fewer cross connections [links]” (Burrows & Mooring, 2015, p. 53). The novice map will have less content and fewer links in their maps, whereas, the expert map will be more web-like and have more concepts and more links as represented in Abd-El-Khalick and Lederman (2000). The expert map will also have more instances of branching, merging, and interconnectedness. These aspects of concept mapping exhibit a more relational and deep understanding about how the author connects the concepts to various concepts. Markham et al., (1994) studied a group of non-biology major
students and compared them with biology major students. They found similar trends among the major biology students (assumed experts) and non-major biology (assumed novices) in the complexity and content of their maps.

Researchers and teachers would agree that the knowledge structure of an expert versus novice would vary in content and structure. One way that researchers have studied expert/novice differences are comparing majors versus non-major courses (Markham et al., 1994). While researching these courses and the coursework, several differences in course expectations and course content may drive the differences of content delivery and knowledge acquisition. Specifically, just because a student may be in one class or another, they may have a knowledge structure like an expert when they are in a science non-major class or a student may have a novice knowledge structure when in a science major’s course. These differences are not well documented. Thus, more research would be necessary to dissect such differences and comparisons among expert/novice concept maps.

In addition, more research is necessary to review how an expert arranges or organizes the knowledge for their concept map. Researchers have published an expert map, but provided little detail in regards to how their map differs among other experts in other fields of science or education (Abd-El-Khalick & Lederman, 2000). Reviewing several expert maps in a content area has not been published and may provide useful information to researchers and teachers alike. Specifically, this information may be useful in studying conceptual understanding of a content area over time. Equally, studying novice maps may provide key information about content and connections that are retained throughout the student’s coursework. Also, monitoring the student’s conceptual understanding as they progress towards the content of an expert throughout
their coursework would be a type of longitudinal study that would be important for any college program to review for updating curriculum.

Expert versus novice structural knowledge represented in concept mapping has been studied by researchers (Burrows & Mooring, 2015; Markham et al., 1994). However, researchers that have published the results have provided somewhat limited information regarding the expert’s background and why they would be classified as an expert. Just as collecting several artifacts or assessments is necessary for a classroom, there have been limited published reflecting several maps from various experts. Sometimes the teacher of the classroom is considered the expert and compares the student maps. While capturing the teacher concept map can be informative regarding the teaching of the content, there may be missing concepts in the knowledge structure. In addition, there has been limited research comparing the teaching content expert versus a scientific researcher expert. Furthermore, the challenge of time during implementation is a relevant problem. More research is necessary to observe what changes occur over what amount of time for students who are novice and then become expert. More work is necessary to observe how this process occurs during learning and what impact this can have upon the scientific teaching or research field.

Summary

The purpose of this literature review was to review the following three research areas: mapping knowledge structures, summative and formative assessment in a science classroom and concept mapping as formative assessment in a science classroom. Many sources of literature have been used to culminate the diversity of literature that has been used to establish arguments for each of the three research areas. However, more research is necessary to implement all three aspects research topics successfully in a science classroom.
Mapping knowledge structures can provide information of how knowledge is organized. Concept mapping in a classroom requires instruction and practice. Much of the literature provides limited information of how the students practice concept mapping and how they were instructed to complete the concept mapping exercise. Providing students the opportunity to train and practice concept mapping prior to its formal use may provide useful feedback to the instructor regarding content delivery, knowledge gaps and misconceptions. The opportunity to utilize the same assessment throughout the semester with little variation may be useful in gathering serial information of learning progressions about a particular topic area of interest. However, serial assessments take time to plan and classroom time to implement that is not always available. Often, the high-stakes testing and summative tests that fulfil the final grade that students require for a class takes precedence over the opportunities to review how a student thinks about the content or gaining feedback of content delivery, knowledge gaps, and misconceptions.

Concept mapping can be a versatile and insightful way to distinguish key ideas about a specific topic in a science classroom. Many topics within science have many relationships and connections to other contexts and coursework. More research is necessary to provide valid and useful resources to science educators that align with how knowledge can be structured and related in teaching science.

Research Questions

Using the literature reviewed, there are gaps that have been identified in the literature. This study will attempt to address those gaps by answering the following questions:

- What impact does concept map training have upon the complexity of students’ concept maps?
• What changes do students exhibit in their concept map characteristics throughout a 16-week science course?

• What impact does concept mapping assessment design have upon students’ concept maps?
  o What differences exist between less-directed and more-directed approaches to concept mapping?
  o What impact does the focus question have on a concept mapping task in terms of the knowledge complexity and content of the resultant concept map?

• What differences exist between students with high scoring concept maps and students with low scoring concept maps in their structural knowledge, as represented by concept maps?

• What differences and similarities can be observed between student concept maps intended to assess structural knowledge and student responses to an open-ended assessment intended to assess content knowledge?
CHAPTER III: METHODOLOGY

Class and Sample Description

The course that was used in this study was an undergraduate biology class, “Introduction to the Genetic and Molecular Basis of Life” at a midsized Midwestern university. Each section of the course enrolls up to 24 students each semester. Most of the students taking this course pursue elementary teaching degrees and are therefore pre-service elementary teachers. Typically, the class enrollees are young, Caucasian, female students. The course was taught as a blended lab/lecture format, meeting two times per week for 2 hours and 40 minutes each meeting. Topics for this course covered aspects of nature of science and cell biology. The course curriculum covered four basic units: Nature of Science (NOS), Cells, Cell Division and Genetics, and DNA and Biotechnology. Each unit had pre-determined learning objectives that aligned with national and state standards while covering the necessary benchmarks of life science for pre-service elementary teachers. Because of the explicit and reflective integration of NOS throughout the course and within most of the engaging class activities, the researcher expected students to develop a more sophisticated understanding of NOS through the semester. Therefore, nature of science was used as the content focus for this study.

Data Sources and Timeline

A total of twelve concept maps and two Views of Nature of Science (VNOS) instruments were collected throughout the Spring 2012 semester. Table 1 delineates the differences of assessment using the focus question of the map (where applicable), the content focus, and the assessment design. Each assessment had either a narrow or wide content focus. Narrowly focused content was considered a specific unit of the course, while widely focused content was
considered as a broad topic that was associated as a learning outcome of the course. All the maps were purely done for research purposes.

Also, the concept map tasks were categorized as directed or less-directed. Novak (1998) used similar definitions by stating maps could be fixed-form or free-form. A directed (fixed form) map will provide either the concepts, link phrases, or the structure for students to utilize in their map. Trowbridge and Wandersee (1996) suggest directed maps are also considered seeded maps when concepts are provided. For this research, the terms directed and less-directed were used as the terms suggest a continuum or variety of task requirements (Table 1).

Concept Maps

The first concept map was collected prior to any instruction and had the focus question, “What is science?” (Map 1, Table 1). The purpose of this map was to determine what students already knew about how to represent their own knowledge structure through the use of concept mapping.

Table 1  
Assessments implemented per research question

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Day</th>
<th>Content Focus</th>
<th>Assessment Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Question 1: What impact does concept map training have upon the complexity of students’ concept maps?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is science? (Pre-Training Map, Map 1)</td>
<td>1</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Post-training Map; Map 2)</td>
<td>2</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>Research Question 2: What changes do students exhibit in their concept map characteristics throughout a 16-week science course?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is science? (Map 2)</td>
<td>2</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 3)</td>
<td>3</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 4)</td>
<td>4</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 5)</td>
<td>8</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 6)</td>
<td>16</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 7)</td>
<td>21</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 8)</td>
<td>30</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
</tbody>
</table>
Table 1, continued

<table>
<thead>
<tr>
<th>Assessments</th>
<th>Day</th>
<th>Content Focus</th>
<th>Assessment Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Question 3: What impact does concept mapping assessment design have upon students’ concept maps?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Question 3a: What differences exist between less-directed and more-directed approaches to concept mapping?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOS Quiz</td>
<td>5</td>
<td>Narrow</td>
<td>No Focus Question; Directed Map</td>
</tr>
<tr>
<td>Mitosis Map</td>
<td>10</td>
<td>Narrow</td>
<td>No Focus Question; Directed Map</td>
</tr>
<tr>
<td>Research Question 3a: What differences exist between less-directed and more-directed approaches to concept mapping?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is science? (Map 5)</td>
<td>8</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>Mitosis Map</td>
<td>10</td>
<td>Narrow</td>
<td>No Focus Question; Directed Map</td>
</tr>
<tr>
<td>What is science? (Map 6)</td>
<td>16</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>Research Question 3b: What impact does the focus question have on a concept mapping task in terms of the knowledge complexity and content of the resultant concept map?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is science? (Map 4)</td>
<td>4</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>NOS Quiz</td>
<td>5</td>
<td>Narrow</td>
<td>Less-Directed Map</td>
</tr>
<tr>
<td>What is science? (Map 5)</td>
<td>8</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>Research Question 3b: What impact does the focus question have on a concept mapping task in terms of the knowledge complexity and content of the resultant concept map?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is biotechnology?</td>
<td>29</td>
<td>Narrow</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 8)</td>
<td>30</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>Research Question 4: What differences exist between students with high scoring concept maps and students with low scoring concept maps in their structural knowledge, as represented by concept maps?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is science? (Pre-Training Map, Map 1)</td>
<td>1</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 2)</td>
<td>2</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 3)</td>
<td>3</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 4)</td>
<td>4</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>NOS Quiz</td>
<td>5</td>
<td>Narrow</td>
<td>No Focus Question; Directed Map</td>
</tr>
<tr>
<td>What is science? (Map 5)</td>
<td>8</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>Mitosis Map</td>
<td>10</td>
<td>Narrow</td>
<td>No Focus Question; Directed Map</td>
</tr>
<tr>
<td>What is science? (Map 6)</td>
<td>16</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 7)</td>
<td>21</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is biotechnology?</td>
<td>29</td>
<td>Narrow</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 8)</td>
<td>30</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
</tbody>
</table>
Table 1, continued

Assessments implemented per research question

<table>
<thead>
<tr>
<th>Assessments</th>
<th>Day</th>
<th>Content Focus</th>
<th>Assessment Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Question 5: What differences and similarities can be observed between student concept maps intended to assess structural knowledge and student responses to an open-ended assessment intended to assess content knowledge?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is science? (Map 2)</td>
<td>2</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 3)</td>
<td>3</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 4)</td>
<td>4</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 5)</td>
<td>8</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 6)</td>
<td>16</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 7)</td>
<td>21</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>What is science? (Map 8)</td>
<td>30</td>
<td>Wide</td>
<td>Less-directed Map</td>
</tr>
<tr>
<td>Pre-instruction VNOS</td>
<td>1</td>
<td>Narrow</td>
<td>Open Ended</td>
</tr>
<tr>
<td>Post-instruction VNOS</td>
<td>30</td>
<td>Narrow</td>
<td>Open Ended</td>
</tr>
</tbody>
</table>

The next class (Day 2), students received explicit training, by the researcher, about why concept maps are useful, how to create concept maps, the components of a useful concept map, and what “good” concept maps look like (Anohina-Naumeca, 2015; Burrows & Mooring, 2015; Eppler, 2006; Quinn et al., 2003-04). Students were trained to use a network/wheel approach in creating their maps instead of a simpler linear, circular, tree, or hub-spoke shape (Meagher, 2009; Yin et al., 2005). The 45-minute training included the researcher teaching and provided students an opportunity to practice the concept mapping. Students practiced creating an individual concept map that answered a non-science focus question, “What is vacation?” This interactive training guided students in creating a list of their own concepts about vacations that they then related to each other using arrows and linking words. They were also instructed and practiced how to include links and linking phrases that created branching and merging occurrences. Students created this map individually, but then they shared their maps with peers for feedback. After receiving peer feedback, students were encouraged to update their own maps. Because peer feedback and modification of the maps occurred on this map, this map was
excluded from analysis. On Day 2, after the training and creation of the vacation practice map, students were asked to create their second “What is science?” concept map (Map 2). This map is considered their first, “What is science?” post-training map. Whenever a mapping task was assigned throughout the semester, each student would individually create a new concept map without peer or instructor feedback. The maps were never returned to the student for review or self-reflection.

On Day 3 and Day 4, the students completed two more “What is science?” maps (Map 3 and 4). These maps were completed near the end of the day after students had been through an early NOS unit that was taught using an explicit and reflective approach. On Day 5, students completed the NOS quiz, which asked students to identify one of the four main NOS activities they performed in the NOS unit: Tube, Twirlie, Fossils, and Cube. Like other mapping tasks, students had to create a list of concepts, rank the concepts, and map the concepts relating to that activity. The directions were the same on the NOS quiz as the “What is science?” maps, which suggested to the student to be cognizant of including links, linking phrases, levels of hierarchy, branching, and merging. This quiz was counted as credit or no credit for their participation. The NOS quiz therefore had a narrow content focus, no focus question, but was a less-directed map. The fifth “What is science?” map (Map 5) was implemented on the day the NOS unit ended. This map was completed with the same parameters as Maps 3 and 4.

On day 10, students created a directed concept map on the subject of mitosis. The instructor provided eleven terms in the map instructions regarding mitosis: mitosis, chromosome, chromatid, interphase, cell cycle, synthesis, metaphase, centromere, spindle fibers, daughter cells, and centrosome. The terms provided were selected as common words from the students’ course pack and the learning objectives from the unit. The purpose of this assessment was to
determine how students would represent their knowledge given these terms using this directed concept mapping approach to a narrow topic focus. Providing terms to a concept map is considered a more directed approach than the other concept mapping tasks when students provide their own concepts, linking phrases, and design their own structure (Brandstädter et al., 2012; Yin et al., 2005). By comparing multiple assessment designs, a topic outside of NOS was necessary to choose for this directed map because giving NOS terms could have influenced their word choices for the future “What is science?” maps. Two additional “What is science?” maps (Map 6 and Map 7) were completed after the mitosis map, on Days 16 and 21.

On Day 29, students completed the presentation of their biotechnology projects and were asked to complete a concept map answering the focus question, “What is biotechnology?” This map was considered a less-directed map which followed a similar format to the “What is science?” maps. However, this map had a narrow focus question. The reason the biotechnology topic and focus question were chosen was because biotechnology inherently incorporated many of the aspects of NOS that students had been exposed to all semester. However, unlike the content in the other units, the instructor rarely explicitly discussed NOS in terms of biotechnology. The researcher wondered if students would transfer their knowledge of science in general to the topic of biotechnology, or if narrowing the focus question to biotechnology would produce a different knowledge structure for students. On the final day of instruction, the students were asked to complete their last “What is science?” concept map (Map 8). The concept maps collected for this research were used to understand how students related the concepts of the course and not used as formative or summative assessment.
VNOS Instrument

The VNOS instrument was implemented on the first day of the class prior to any content instruction. Students completed the VNOS after their pre-training concept map and were expected to answer the questions to the best of their abilities. The purpose of the VNOS instrument was to provide students an opportunity to respond to a narrowly focused, open-ended instrument that represented a type of assessment in this research. On the last day of class, the VNOS was completed by students again after the final concept mapping activity and was used as a post-instruction instrument.

Interviews

Students were asked to participate in a post-instruction interview during the last week of the semester and researchers sought about 10% of the consented population. Unfortunately, three students agreed to participate in the interview, and only one student showed up to complete the interview using their completed concept maps and VNOS instruments. By only having one participant, the analysis of the interview was not used in the data analysis.

Data Analyses

Concept Map Analysis

Concept maps created by the students were analyzed by counting the following concept mapping structural characteristics: total concepts, total links, levels of hierarchy, occurrences of branching, and occurrences of merging (cross-linking). The total number of concepts were counted as the total number of unique terms, identified by the student, that are linked together with lines and possible phrases. The total concepts contained a subgroup, NOS concepts, which counted as the following concepts: creative, subjective, tentative, theory, law, empirical, inferences, scientific method (multiple methods), social & cultural embeddedness, social
dimensions, and models. The NOS concepts value was included in the total concepts value. The modeling term while not related to NOS directly was grouped in this subgroup for this research study.

The total links were counted in this study and not the propositions because students would not always include linking phrases in their maps. Hierarchy was counted as the number of steps or levels in the longest chain not including the top concept (Novak & Gowin, 1984). Branching was defined as a single concept delineating into two or more concepts (Markham et al., 1994; Wallace & Mintzes, 1990). Whereas, Afamasaga-Fuata’I and Reading (2007) defined merging as two or more concepts are combined into a single concept. However, cross-linking, which is the linking of non-linear concepts, is essentially the same thing as merging (Markham et al., 1994; Martin et al., 2000; Quinn et al., 2003-04). Thus, for this research when two concepts are combined into one, the occurrences were counted as merging. An example of a scored map is provided (Figure 1). This map (#169) was coded having 14 total concepts (squared), 17 links (lines), eight levels of hierarchy (diamonds with numbers), four occurrences of branching (green circles), and four occurrences of merging (yellow triangles). The averages were calculated for each mapping event. In addition, changes between mapping events were compared using percent intervals changes of the mean to determine what structural components increased or decreased.

A raw score was calculated by adding together the total concepts, total links, levels of hierarchy, occurrences of branching, and occurrences of merging. This method was chosen for this research because researchers have not agreed upon how each structural component should be multiplied by numerical variables to evaluate its influence (Novak & Gowin, 1984; Markham et al., 1994; Quinn et al., 2003-04). The calculated raw score from this study provided an unweighted individual score for each map.
Using these individual scores, the concept maps were ranked into terciles to determine what students were considered “experts” and what students were considered “novices.” Terciles were used to calculate the three levels of scorers, low-scorers, middle-scorers, and high-scorers (Burrows & Mooring, 2015).

The overall concept map shape was analyzed using previously published shapes (Kinchin, Hay, & Adams, 2000; Meagher, 2009; Yin et al., 2005). Overall structural comparisons were used to classify each map into five categories: linear, circular, tree, hub and spoke, network/wheel shapes. By using these pre-designed structures from the literature, the overall shapes provided correlations to a more or less sophisticated understanding of the topic. Figure 2 is an example of a hub/spoke structure.
However, when inter-coding was completed on a sub-sample of the total number of maps, only a 75% agreement was met and it was decided that more distinguishing factors were necessary to code the post-training maps beyond using the pictures that the literature provided (Kinchin et al., 2000; Meagher, 2009; Yin et al., 2005). Thus, more work is necessary to determine what quantitative factors or other qualitative factors would clearly distinguish differences in the overall map structure.

Statistical Analyses

A two-sample separate variance t-tests was used to test for differences between the pre-training and post-training maps in their structural complexity. Additionally, a two-sample separate variance t-tests were used to compare differences in the structural complexities between a less-directed NOS quiz and directed Mitosis map. Two-sample separate variance t-tests were used to test for the difference in structural complexities between two different focus question maps and the difference in structural complexities between the low-scorers to high-scorers concept maps.
VNOS Analysis

Each VNOS instrument was blinded, reviewed, and analyzed using the coding procedure adapted from McDonald (2008). However, the categories used in this research were updated to reflect categories that easily distinguishable. The researcher was trained on how to evaluate the students’ responses on the VNOS. Each of the NOS tenets (Creative, Subjective (Theory-laden), Tentative, Theory/Law, Empirical, Inferences, Myth of Scientific Method, Social & Cultural Embeddedness, Social Dimensions) were analyzed by coding each tenet using the following five categories: generally incorrect views of that tenet, confused views of that tenet, correct but simplistic views of that tenet, consistently correct views of that tenet, and consistently correct views of that tenet with the ability to provide examples of it from practice. Frequencies of each category were calculated for the 17 participants, for both pre- and post-assessments, for each of the NOS tenets. Inter-coder reliability was established by coding 10 percent or four of the 34 VNOS instruments.

Confidentiality

All the concept maps and VNOS instruments were collected by the instructor and held for the researcher until the end of the semester. The instructor did not return the concept maps to the students once the maps were completed by the students. The concept maps were randomly coded by another research and prior to analysis. Randomly coded concept maps were double blinded in regards of the participants name and date of the concept maps being completed.

Role of the Researcher

The researcher had previously taught this class during a summer semester and co-taught the class during a 16-week fall semester. Thus, the researcher was familiar with the course content, course objectives and expectations, as she helped create and modify the assignments and
some of the content. In this study, the researcher held only the role of an observer of the course and reviewed and analyzed the students’ work after the semester ended. By completing this methodology, the researcher hoped to gain insight into being able to replicate the role of concept maps in other courses and domains of science.

Limitations to the Study

The research was limited by the number of participants that were willing to participate in the research, given the small class size and limited opportunity to collect the data (restricted to a single semester). To remediate this as much as possible, multiple data sources were used per student, which helped triangulate any findings and provided enough numerical data to be able to form useful conclusions about the role of concept mapping as an assessment. In addition, it is difficult to know if the students were engaging in the NOS activities during the class time. The researcher assumed that the students who attended were all engaged in the explicit nature of science activities.

Summary

This study used multiple concept map assessments and two VNOS instruments that were completed by students to monitor how their knowledge structure changed throughout a semester. The concept maps were analyzed in their overall shape and their structure complexity. Raw scores calculated from the structure complexity and were used in calculating terciles for comparing expert to novice maps. In addition, the VNOS was coded by using five categories to describe the overall class’ frequencies of each NOS tenet. The confidentiality of the assessments was maintained and the researcher held an observational role in the classroom. Other researchers were useful in blinding the data to maintain confidentiality.
CHAPTER IV: RESULTS

Introduction

The purpose of this chapter is to review the description of the sampled participants and examine the data as it relates to the research questions. Major trends, differences, and similarities were used to highlight key information of all the knowledge structures obtained through student concept mapping and student responses to a concept-assessment for nature of science (Views of Nature of Science - VNOS). This chapter will provide evidence for conclusions using figures and tables, so the data can be discussed and related to the overall theoretical framework of the study.

Description of the Sample

The research was conducted at a Midwestern university in an undergraduate biology class. All student participants were enrolled in the “Introduction to the Genetic and Molecular Basis of Life” course and completed a pre-instruction VNOS, post-instruction VNOS, and twelve concept maps throughout the 16-week semester. A total of 17 undergraduate students (N=17) agreed to participate in the research study, including 16 females and one male. For two of the participants, this was their only science course taken so far at the college level. The other 15 participants were either taking a co-concurrent science class or had previously taken science courses at the college level. No other self-identifying descriptors were collected from the participants. Student participation was voluntary and informed consent was obtained. To protect their identity, their names were replaced with random numerical codes.

Details of the Analysis

Research Question 1: What impact does concept map training have upon the complexity of students’ concept maps?
The pre-training maps and post-training maps were analyzed for their overall structural shape and their structural components (Afamasaga-Fuata’I & Reading, 2007; Markham et al., 1994; Martin et al., 2000; Novak & Gowin, 1984; Quinn et al., 2003-04; Wallace & Mintzes, 1990). The knowledge structure complexity was measured as five components, total concepts, total links, hierarchy, branching, and merging. All five components were significantly higher in the post-training concept maps ($p < 0.05$, Figure 3).

![Figure 3. Effects of concept map training on structural component means ± SD (n=34).](image)

Merging (4,500%) and Total Links (3,329%) had the greatest percent change between the pre-training and post-training maps. When comparing all five structural components each time, total concepts had the highest mean, while merging had the lowest mean (Figure 3). When comparing the overall structural shape, the hub/spoke structure was the dominant structure with a frequency of 70.6 of all the pre-training maps (Figure 4). Linear maps were not observed in either mapping event. Both the network/wheel and tree structural shape increased after the
concept map training (Figure 4). The post-training map was also included as “What is science?” post-training analyses and labeled as Map 2.

Figure 4. Comparison of pre-training and post-training structural shape frequencies (n=34).

Research Question 2: What changes do students exhibit in their concept map characteristics throughout a 16-week science course?

All the post-training “What is science?” maps were compared using the five knowledge structural components. The maps were compared by grouping the structural components together to determine what trends may exist (Figure 5). The largest mean for merging occurred on the post-training map (“What is science?” Map 2) as seen in Figure 5, showing a decreasing trend through the remaining “What is science?” maps the rest of the semester. Concept total reached its highest mean on Map 5, which is also the highest mean for total links and hierarchy (Figure 5). The “What is science?” post-training maps were also compared by structural components using
percent changes between mapping events (intervals). The percent interval was calculated by taking the difference between the structural component means (Figure 6).

Figure 5. Comparing structural components using the same focus question for post-training. “What is science?” Map 2 through 8. Values are means ± SD (n=114).

For post-training “What is science?” Map 2 and Map 3, the hierarchy was the only structural component that increased while all the other components decreased (Figure 6). Total concepts increased from Map 3 to Map 4, however all the other structural components percent interval decreased between those two maps. Whereas, all the structural components increased between the Map 4 to Map 5 interval and the Map 7 to Map 8 interval (Figure 6). In both intervals, merging had the greatest percent change compared to the other mapping components.

Research Question 3a: What impact does concept mapping assessment design have upon students’ concept maps? Specifically, what differences exist between less-directed and more-directed approaches to concept mapping?
Figure 6. Percent changes between “What is science?” post-training concept maps (n=114).

On the NOS quiz, students were asked to list concepts that they learned from one of the four NOS activities provided on the quiz (Twirlie, Fossil, Tube and Cube), all of which were done in class. They were to rank their listed concepts and map the concepts. This map was considered a less-directed map because students had to generate their own concepts, links, and map structure, but the content had a narrow focus. The quiz was not graded for correctness but only for attempting the mapping activity. The mitosis map was considered a narrow content map as well since the topic was specific to mitosis. However, because eleven terms were given to the student, the mitosis map was considered a directed map. Merging was the only structural components on the NOS quiz (less-directed) maps that was significantly different than those on the Mitosis (directed) maps \( (p < 0.05) \). This significance is despite having more total concepts, total links, and branching occurrences in the NOS quiz (Figure 7). The Mitosis (directed) map
had more levels of hierarchy identified, but this difference was not significantly greater than the NOS quiz. However, merging was significantly greater in the NOS quiz ($p < 0.05$, Figure 7).

![Figure 7](image)

Figure 7. Comparison of structural components of a less-directed to a directed map. Values are means ± SD (n=32).

The “What is science?” Map 5 and Map 6 were compared to the Mitosis (directed) map. The purpose of this comparison was to observe what changes may occur when giving students other mapping assessments compared to their serial, “What is science?” maps that were created throughout the semester. Map 5 was completed prior to the Mitosis (directed) map and Map 6 was the next map students completed following the Mitosis (directed) map (Table 1). For the Mitosis (directed) map, students were asked to map the provided eleven concepts about mitosis. This approach asked students to use the eleven concepts but did not say they could not use additional concepts. Whereas, students were asked to generate their own list of concepts, rank the concepts, and map the concepts for the “What is science?” maps.
Observing the trends among “What is science?” maps and the Mitosis (directed) map indicate that the narrowed-content focus, directed map showed a lower mean for total concepts, total links, branching, and merging (Figure 8). However, the hierarchy of the Mitosis (directed) map increased in comparison to the wide content focus, less-directed, “What is science?” maps (Figure 8).

Figure 8. Comparing the integration of a directed mapping task among serial maps. Values are means ± SD (n=44).

Research Question 3b: What impact does concept mapping assessment design have upon students’ concept maps? Specifically, what impact does the focus question have on a concept mapping task in terms of the knowledge complexity and content of the resultant concept map?

Knowledge structure complexity was also compared between two types of less-directed approaches using “What is science?” maps and the NOS quiz. The “What is science?” maps provided students a wide content focus questions to answer and was a less-directed assessment design. The NOS quiz used a narrow content focus but was also a less-directed assessment design. The NOS quiz was completed between the two “What is science?” maps on their
respective days (Table 1). Another purpose of the quiz was to determine what concepts students identified from their knowledge when prompted to associate it with prior activities and if their knowledge was represented differently from their “What is science?” map. The quiz had fewer total concepts than the two “What is science?” maps, but similar total links from Map 4 to NOS quiz (Figure 9). The hierarchy gradually increased between all three maps. Branching decreased in the NOS quiz between the two “What is science?” maps, but merging increased from Map 4 to the NOS quiz and had the same average between the NOS quiz and Map 5 (Figure 9).

Figure 9. Comparing two less-directed assessment designs, one with a wide focus question and another with a narrow content focus. Values are means ± SD (n=50)

The diversity of NOS concepts was analyzed and compared among all three maps. These three maps were used to determine what trends exist in student knowledge of NOS content through this time segment. The total concepts decreased for the NOS quiz compared to both “What is science?” maps (Figure 9). However, the NOS quiz increased in the total frequency of NOS terms despite having fewer total concepts (Figure 10). Therefore, within fewer total
concepts, the students included more NOS concepts, compared to the “What is Science?” maps. The concept, Scientific Method/Multiple Methods, was not integrated by students in their “What is science?” Map 4, whereas in the NOS quiz and “What is science?” Map 5, the concept of Social Dimensions was not integrated (Figure 10).

Figure 10. Comparing total frequencies and diversity of NOS concepts among two types of less-directed assessment designs (n=50).

Two concept mapping tasks were compared using different focus questions, but with related content. The two focus questions of the maps were “What is biotechnology?” and “What is science?” (Map 8, Figure 11). The biotechnology map was completed on the last day of presentations and prior to the final exam week, which was the day before the final “What is science?” map (Map 8, Table 1). The biotechnology map used a focus question relating to the final unit that covered issues related to biotechnology, such as Genetically Modified Organisms (GMOs). The “What is science?” Map 8 had a greater mean of total concepts (Figure 11), but the
differences between each of the structural components for both maps were not significant. The “What is science?” Map 8 had more levels of hierarchy and occurrences of branching, but similar averages of merging occurrences (Figure 11).

Figure 11. Comparing structural components using different focus questions. Values are means ± SD (n=33).

Furthermore, the “What is science?” Map 8 had greater total frequency of reported NOS concepts and a greater diversity of NOS concepts than the “What is biotechnology?” map (Figure 12). The “What is biotechnology?” map had three NOS concepts identified in the maps, Creative, Subjective, and Social & Cultural. Whereas, the “What is science?” Map 8 included all NOS concepts but the Social Dimensions concept (Figure 12). Neither of the maps identified the NOS concept, Social Dimension, in their maps.
Research Question 4: What differences exist between students with high scoring concept maps and students with low scoring concept maps in their structural knowledge, as represented by concept maps?

Figure 12. Comparing NOS concepts between two less-directed, but different focus question maps (n=33).

All eleven concept map assessments were sorted by their raw score (n=120) to distinguish the terciles of scores for novice and expert maps. The raw score of a map was calculated as the sum of the total concepts, total links, hierarchy, branching, and merging that occurred in the map. Terciles were used to distinguish low-scorers (Novice) to high scorers (Expert) for each mean of the mapping structural components. For this research, middle scorers were not used in the analysis as compared in the expert/novice concept mapping literature (Burrows & Mooring, 2015). When comparing the low-scorers (Novice) versus the high-scorers (Experts), all the structural components were significantly higher for the high-scorers ($p < 0.001$, Figure 13). The frequency of NOS concepts was also compared between the low-scorers and
high-scorers (Figure 14). NOS concepts of high-scorers had a greater frequency than low-scorers with one exception, the NOS concept, Inferences (Figure 14).

Figure 13. Significant differences exist between all structural components between low and high scorers ($p < 0.001$). Values are means ± SD (n=120).

Figure 14. Comparing NOS concepts between low-scorers and high-scorers (n=120).
Inference was mentioned more frequently in the low-scorer maps than high-scorer maps. Three NOS concepts, Creative, Subjective, and Theory, had the greatest frequency in the high-scorer maps. But, the Social Dimension concept was the lowest frequency for all the high-scorer maps (Figure 14).

Research Question 5: What differences and similarities can be observed between student concept maps intended to assess structural knowledge and student responses to an open-ended assessment intended to assess content knowledge?

NOS concepts were identified in the post-training, “What is science?” maps (Figure 14). The purpose of identifying these concepts was to monitor what content changes occurred throughout the student knowledge structures using the same focus question on concept maps throughout the semester (Table 1). For this analysis, only the post-training concept maps were used, as it was determined that training had a significant impact on map quality. However, there was not one mapping event that contained all the NOS concepts. Subjective and Theory concepts had the greatest total frequency for all the “What is science?” maps (Figure 15).

Again, the Social Dimensions concept had the lowest total frequency of occurrences and was represented in the fewest post-training, “What is science?” maps. When comparing the total frequencies of all the NOS concepts, Map 6 had the greatest total frequency with 492 occurrences. Whereas, when including the modeling concept with NOS concepts, the greatest total frequency was 523 for both Map 8 and Map 6. The “What is science?” Maps 5-8 were only missing the Social Dimensions concept, but contained all the other NOS concepts (Figure 15).

In addition, a pre-instruction and post-instruction VNOS was used to compare short-answer, written responses to the knowledge structures of students as assessed by their concept maps (Table 2 and Table 3, respectively). The purpose of the VNOS was to create a NOS profile
of the class prior to teaching NOS tenets and at the end of the class to determine what NOS tenets were lost, retained, or improved upon in a post-instruction profile identified in the VNOS.

Figure 15. NOS concepts identified in post-training “What is science?” maps (n=114).

Creative, Subjective, and Tentative are the NOS tenets that increased throughout the semester in both the concept maps and the VNOS (Figure 15, Table 2, Table 3). Scientific method/multiple methods concept was not identified in the concept maps but showed a positive shift in understanding from the VNOS (Figure 15, Table 3). However, there was an increase in the concept, Models, identified in the concept maps over time (Figure 15). Modeling concept on the VNOS is often associated with the tenet of Multiple Methods/Myth of Scientific Method, and this tenet shifted to a more consistently correct tenet (Table 3). The NOS concept, Empirical, was used more in the final “What is science?” map (Figure 15), and students also described science as being evidence-based more consistently in their post-instruction VNOS (Table 3). The students had fewer incorrect descriptions of the NOS tenet, Inferences, on their post-instruction
VNOS, but were still confused (Table 3). Interestingly, Inferences was the only NOS concept the low-scorers identified more than did the high-scorers on their concept maps (Figure 14).

Table 2
*Pre-instruction VNOS analysis comparing NOS views per tenet*

<table>
<thead>
<tr>
<th>NOS Tenet</th>
<th>% Incorrect</th>
<th>% Confused</th>
<th>% Correct</th>
<th>% Consistently correct</th>
<th>% Consistently correct with example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creative</td>
<td>5.88</td>
<td>35.29</td>
<td>35.29</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Subjective</td>
<td>17.65</td>
<td>17.65</td>
<td>47.06</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tentative</td>
<td>58.82</td>
<td>29.41</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Theory/Law</td>
<td>85.29</td>
<td>11.76</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Empirical</td>
<td>17.65</td>
<td>41.18</td>
<td>23.53</td>
<td>11.76</td>
<td>-</td>
</tr>
<tr>
<td>Inferences</td>
<td>29.41</td>
<td>58.82</td>
<td>5.88</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Myth of Scientific Method</td>
<td>29.41</td>
<td>52.94</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Social &amp; Cultural</td>
<td>-</td>
<td>5.88</td>
<td>5.88</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Social Dimensions</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*Note.* Percentages of NOS views do not always add up to 100% because not all participant responses were coded for these tenets.

Summary

Comparing the impact of concept map training, serial concept mapping, assessment design, low-scorers to high-scorers, and reviewing the concept inventory VNOS to concept mapping was done for this data analysis. A total of 179 concept maps and 34 VNOS assessments were collected and used for this analysis. The assessments varied by implementation timing, less-directed or directed assessment design, and wide or narrow-content focus of assessment. A variety of assessments were used to better understand how multiple assessments can be used in a science college level classroom to monitor knowledge structures throughout a semester.
Table 3
*Post-instruction VNOS analysis comparing NOS views per tenet*

<table>
<thead>
<tr>
<th>NOS Tenets</th>
<th>% Incorrect</th>
<th>% Confused</th>
<th>% Correct</th>
<th>% Consistently correct</th>
<th>% Consistently correct with example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creative</td>
<td>-</td>
<td>5.88</td>
<td>52.94</td>
<td>35.29</td>
<td>5.88</td>
</tr>
<tr>
<td>Subjective</td>
<td>-</td>
<td>5.88</td>
<td>35.29</td>
<td>11.76</td>
<td>47.06</td>
</tr>
<tr>
<td>Tentative</td>
<td>11.76</td>
<td>35.29</td>
<td>29.41</td>
<td>17.65</td>
<td>-</td>
</tr>
<tr>
<td>Theory/Law</td>
<td>55.88</td>
<td>14.71</td>
<td>11.76</td>
<td>5.88</td>
<td>-</td>
</tr>
<tr>
<td>Empirical</td>
<td>-</td>
<td>-</td>
<td>11.76</td>
<td>52.94</td>
<td>23.53</td>
</tr>
<tr>
<td>Inferences</td>
<td>11.76</td>
<td>41.18</td>
<td>23.53</td>
<td>5.88</td>
<td>5.88</td>
</tr>
<tr>
<td>Myth of Scientific</td>
<td>11.76</td>
<td>41.18</td>
<td>17.65</td>
<td>29.41</td>
<td>-</td>
</tr>
<tr>
<td>Method</td>
<td>Social &amp; Cultural</td>
<td>11.76</td>
<td>17.65</td>
<td>5.88</td>
<td>-</td>
</tr>
<tr>
<td>Social Dimensions</td>
<td>5.88</td>
<td>11.76</td>
<td>23.53</td>
<td>5.88</td>
<td>-</td>
</tr>
</tbody>
</table>

*Note.* Percentages of NOS views do not always add up to 100% because not all participant responses were coded for these tenets.
CHAPTER 5: DISCUSSION

Introduction

Using concept maps to monitor knowledge changes can help researchers better understand what connections between concepts students are making when learning. Student responses to short answer, essay, or multiple choice assessments do not always articulate those relationships well (Lazarowitz & Lieb, 2006; Markham et al., 1994; Weurlander et al., 2012). This study focused on the mapping assessment itself and how serial mapping and different map designs impact the resultant maps created by students. This study provides an analysis of eleven concept maps that represent students’ knowledge structures and its changes over time.

Discussion of the Results and its Relation to the Literature

Research Question 1: What impact does concept map training have upon the complexity of students’ concept maps?

There was a significant increase in knowledge structure complexity between the pre-training and post-training maps (Figure 3). This would suggest that the concept map training that students participated in was necessary to elicit more knowledge in their structure, which we assume represents conceptual understanding in response to the focus question. This is the first time, per the literature search, that significant changes in pre-training and post-training concept maps have been reported in the literature.

While many studies do mention that researchers conduct concept map training in their methodology (Markham et al., 1994; Martin et al., 2000; Novak & Gowin, 1984; Wallace & Mintzes, 1990), there has been no evidence to suggest what impact the training had upon student concept maps, what students learned from the training, and if the training was effective in
teaching them the skills of concept mapping. This study addresses this small, but important gap in the literature by providing evidence that some students did know how to concept map prior to instruction, but after completing training, the structural components increased and the maps included more total concepts (Figure 3). The significant increase would suggest that training should always be completed prior to the students being assessed of their learning with a concept mapping assessment tool. The impact of having a pre-training and post-training map would create a base-line for researchers or teachers of what mapping skills and structural components should be reinforced in the future.

The overall map shape (as identified by Meagher, 2009) students used in their pre-training maps was a hub/spoke shapes. Post-training maps showed an increase of network/wheel and tree map shapes (Figure 4). Both the network and tree map shapes require more concepts and links which the post-training map contained. The increase of the network/wheel map shape would also support the findings that students included more branching and merging in their post-instruction map (Figure 3). However, the inter-coder analysis of identifying map shapes was only 75%. Even after discussion, there was confusion on how several map shapes should be coded. Thus, more details are required to distinguish clearer differences for this qualitative coding of the overall map shape. A future consideration would be evaluating how many of each of the structural components would be necessary to distinguish the variety of structural shapes proposed by Meagher (2009) and Yin et al., (2005).

Research Question 2: What changes do students exhibit in their concept map characteristics throughout a 16-week science course?

The analysis of the post-training “What is science?” Maps 2-8, which used the same focus question, indicated a decrease in the use of merging throughout the semester. Other small
issues, not directly provided through the data (such as a lack of direction arrows or incomplete linking terms) indicated the need for additional training throughout the semester (Figure 5). Since this study purposefully did not allow students to revisit their previous maps, nor did it provide students with feedback on their maps, the researcher cannot say how feedback or additional training might impact student maps. However, when comparing Map 4 to Map 5, it is unclear why there was a sharp increase in all the mapping components, which then dropped again immediately in Map 6 (Figure 6). The decline of all characteristics from Map 5-7 follows a similar pattern of what occurred prior in Maps 2-4, which may indicate that students forget about mapping components over time, or simply get tired or bored of creating the maps (Figure 6).

Researchers have used serial concept mapping in their college level courses for assessment (Martin et al., 2000; Meagher, 2009; Quinn et al., 2003-04). However, Martin et al., (2000) collected only four maps per semester for two semesters. When asking students to complete the mapping activity, the previous maps were given back to students and students were asked to change or redraw their map. This allowed the student to revisit their previous map and reflect upon what prior knowledge they mapped. Quinn et al., (2003-04) asked students to complete three maps within the semester and another map five to six months after the class completed. Their participants were able to review their previous map and make changes or disregard them. Meagher (2009) asked students to complete three concept maps throughout the semester, but allowed students to focus on any topic of choice. Their choices could vary among the maps they created in one semester.

The research of this study did not allow students to review their previous maps, which was intended to provide a more accurate monitor of the structural and content changes associated with students’ knowledge structures. This methodical design provided an opportunity to review
how concepts and the connections between concepts would endure, or not, throughout the semester (Hay, 2007). Furthermore, when students maintain similar concepts within their knowledge structure, that would suggest that the concepts were integrated into their long-term memory (Meagher, 2009) and were meaningfully learned (Lazarowitz & Lieb, 2006). Also, the changes that occur within the knowledge structure represented in the concept map may suggest what changes occur as a result of instruction (Shavelson et al., 2005).

Having students create their own maps can be useful in monitoring what concepts remain, or not, in their knowledge structure over time. Also, students’ concept map structures did not change throughout the course despite the constant discussion of the concepts as they relate to other topics in the course. This information about retention of concepts will be useful in determine how to address what concepts should be re-addressed in future courses.

In the future, researchers and teachers may want to consider returning the maps back to the students for additional feedback. The revisit of these maps can provide a reflective opportunity as part of formative assessment. This opportunity would give the students a chance to update, redraw or reflect on how their knowledge changed or no. In addition, this may provide a better picture of how concepts can be added to the growing knowledge structure and may see similar trends per Quinn et al., (2003-04).

Research Question 3a: What impact does concept mapping assessment design have upon students’ concept maps? Specifically, what differences exist between less-directed and more-directed approaches to concept mapping?

The assessment design was compared using three comparisons: 1) directed versus less-directed concept mapping assessment design, 2) using two less-directed concept maps, but having one assessment with a focus question and one without, and 3) comparing maps with two
different focus questions but on related topics. A less-directed (NOS quiz) map and directed (Mitosis) map were compared using their structural components (Figure 7). Students were asked to complete a directed mapping activity that provided eleven concepts for them to map about the topic of mitosis. The mitosis map was considered directed because the assessment design limited the total number of concepts, which could have also limited the total links identified in the map. The hierarchy of the Mitosis maps was greater than those on the less-directed (NOS quiz) maps, which is indicative of linear or tree map shape with the fewer instances of branching and merging. A fewer levels of hierarchy is generally found in more complex map shapes, indicating a more complex knowledge structure.

The Mitosis (directed) maps had a lower mean of total concepts, total links, branching, and merging compared to the two less directed maps completed prior and after “What is science?” maps (Figure 8, Table 1). Concepts were provided for the mitosis maps, thus, students mapped what was provided and did not add additional concepts. When fewer concepts are used with fewer links, branching and merging occurrences, this would suggest that simpler maps were used. A more complex map would have an increased number of links, branching and merging occurrences. Thus, when monitoring knowledge structures, a researcher should limit the use of directed map assessment designs. The assessment design constraints of a directed map will limit what students provide as their knowledge structure.

A directed concept mapping assessment design provides some element of the knowledge structure in the assessment. Providing the mapper the concepts, linking phrases (propositions), or the overall structure are all characteristics of a directed map (Ruiz-Primo et al., 2001). Using a directed assessment design to concept mapping can be useful in having students relate key concepts using links and linking terms. This concept map assessment design would not provide
information regarding whether or not the concepts were in the students’ overall knowledge structure. Yin et al. (2005) suggests this particular method would be useful in understanding more about the declarative knowledge students have about a specific topic. Ruiz-Primo et al., (2001) compared a similar approach of less-directed mapping and directed mapping activities. They found that when students created the map with the terms, the students provided more information about each term, which could not be validated in our study. This study compared different content areas to assess differences between directed and less-directed maps, which may contribute to the differences observed in map structure. However, since different content topics were used, the overall representations of understanding between the two maps cannot be compared. This study does though indicate that a directed assessment design does limit the structural complexity of student created concept maps (Figure 8). The overall purpose of the assessment should be evaluated to determine how to assess students’ knowledge structure. Thus, if students are given an assessment several times using the same concepts, this would yield how students are structuring and restructuring those concepts throughout time. This type of assessment may be useful in determining specific relationships among commonly confused or well-known concepts (Burrows & Mooring, 2015). However, if students are given a focus question and the opportunity to re-answer the focus question several times throughout a time, the researchers may see how the overall knowledge structure is being influenced by content or other factors, which was the intent of this research study.

Research Question 3b: What impact does concept mapping assessment design have upon students’ concept maps? Specifically, what impact does the focus question have on a concept mapping task in terms of the knowledge complexity and content of the resultant concept map?
Two comparisons were used to answer this research question. The Nature of Science (NOS) quiz and “What is science?” (Map 4 and 5) were compared to discern any differences that may exist in two less-directed concept mapping assessment designs that address the same concepts. The NOS quiz had students narrow in on a specific NOS activity, while the “What is science?” mapping task was a wide content focus. The NOS quiz was not considered a directed assessment design because the students were still expected to generate all the components of the concept map from their own knowledge and create the map on their own. The NOS quiz had fewer total concepts, which may be a result of the narrow content focus on the question, which referred to a single activity. Both maps had similar averages of total links, hierarchy and branching occurrences. However, merging was higher on the NOS quiz. This would suggest that students created more network/wheel-like shaped maps that utilized merging occurrences to connect non-linear concepts together (Figure 9). Additionally, the NOS quiz had a higher frequency of the common NOS concepts which included modeling (Figure 10). This suggests that the narrow content focus to the NOS quiz allowed students to map more NOS concepts in fewer total concepts in their map. However, after creating the concept map for the NOS quiz, students continued to integrate more NOS concepts in their next “What is science?” (Map 5), but they also had more total concepts as well (Figure 9 and Figure 10).

While comparing two different less-directed concept maps on the same topic, the NOS quiz provides evidence of how students can articulate their increasing amount of NOS concepts in fewer total concepts in their knowledge structures (Figure 9 and Figure 10, respectively). Using both the knowledge structural components and the increasing trend of NOS concepts, concept map quiz objective was supported as a less-directed conceptual focus map. Thus, a narrower content focused map will provide researchers and teachers a smaller scope of content to
focus upon without being limited by the mapping task constraints. This can provide a more focused approach of topic areas that may be useful when completing units of content throughout the semester-long course at the college level.

So, when providing a narrow content focus in the assessment design, this may reduce the number of concepts that students provide. Thus, researchers and teachers should use focus questions that cover a wide content focus to obtain a more representative map of their knowledge structure. However, this research was limited in reviewing relevant connections because of missing linking words/phrases. Therefore, students could have had more meaningful connections that were not observed. Thus, the quality of the relationship was not observed. So, when complexity decreases in quantitative terms, the quality of the concepts might increase. More research is necessary to evaluate the linking phrases in response to the quantitative measurements from this research study. In the future, researchers and teachers could create rubrics that would help set parameters regarding different aspects of the mapping structures.

The second comparison for this research sub-question compares two concept maps with different focus questions, but with similar expected outcomes. These maps were compared to better understand how the question being asked can change the content and complexity of concept maps. On average, “What is biotechnology?” maps had fewer structural components for all five structural components compared to the final “What is science?” (Map 8, Figure 11). However, these differences were not significant and would suggest that students held similar structural complexities between the two topics. The NOS content analysis between the two maps would suggest differently. The “What is science?” Map 8 had a greater frequency of NOS concepts compared to the “What is biotechnology?” map (Figure 12). This difference in NOS concepts would suggest that the students did not integrate the NOS concepts into their conceptual
understanding of the biotechnology unit learning objectives. Students were not able to relate the topics of biotechnology that were researched and presented by their peers to their established NOS conceptual framework or vice versa, despite connections being made between them (Figure 12).

When comparing the concept mapping tool as it relates to the focus question, the knowledge structure complexity was not significant between the two focus questions, “What is biotechnology?” and “What is science?” (Map 8, Figure 11). However, the NOS concepts did vary in their frequency between the two focus questions (Figure 12). Thus, the different focus questions provided an opportunity to review how students applied their biotechnology project that they previously presented to their peers and its impact upon the current NOS understanding. The students included less than 10% of the total frequency of NOS concepts in their biotechnology map (Figure 12). This would suggest that the activities that this curriculum unit correlated with did not focus upon using the explicit, reflective teaching approach of NOS concepts. The evidence from these comparisons in the research study is new to the concept mapping literature, but researchers agree that different mapping techniques yield different results (Shavelson et al., 2005). More work would be necessary to align the biotechnology unit to specific NOS concepts in the future. Furthermore, this comparison highlights again the importance of understanding how the assessment design can impact the results that students provide.

Research Question 4: What differences exist between students with high scoring concept maps and students with low scoring concept maps in their structural knowledge, as represented by concept maps?
Low-scorers (novice) and high-scorers (expert) knowledge structures were compared using structural components and NOS concepts. The low-scorers exhibited a significantly lower mean for each of the structural components than the high scorers \((p < 0.001, \text{Figure 13})\). The high-scorers had a significantly higher mean for each of the structural components which would suggest that their maps were more detailed in concepts and how those concepts are being connected together in a knowledge structure. With more total links, total concepts, and occurrences of branching and merging, the high-scorer maps would be more interconnected among the concepts in the structure. However, the increased hierarchy for high-scorers may suggest that some of the students created structures more in a tree or linear shape that uses branching. Furthermore, the high-scorer maps mentioned NOS concepts more frequently than the low-scorers, except for the concept, Inferences (Figure 14). High-scorers were able to relate NOS concepts in their knowledge structure. The differences in means and inverted relationship may suggest that high-scorers may not have understood the inference concept well enough to identify the concept in their knowledge structure throughout the semester.

Researchers have established that more concepts that relate to each other suggest a more integrated or complex knowledge structure which can be represented in a concept map (Shavelson et al., 2005). Researchers have used various groups of people to compare expert and novice knowledge structures. For instance, Anohina-Naumeca (2015) compared student (novice) to teachers (expert) in their qualitative and quantitative analysis of concept maps (p. 62). Whereas, Burrows and Mooring (2015) used a scoring system of proposition accuracy to then rank student scores into terciles to determine the low-scorers (novice) and high-scorers (expert). All the students were in one organic chemistry class. While Markham et al., (1994) used two college-level courses to compare two student groups, beginning non-majors (novice) and
advanced college biology majors (expert). The researchers compared the two courses in their concept mapping scores. Meanwhile, this study used multiple concept maps from students in an undergraduate course. The raw scores were ranked into terciles and the low-scorers (novice) were compared to the high-scorers (expert, n=120). From these maps, the researcher can conclude that there were significant differences between the novice and expert structural complexities ($p < 0.001$; Figure 13). Markham et al., (1994) observed similar significant differences between the structural components of their maps. Whereas, Burrows and Mooring (2015) observed how high-scoring students meaningfully connected the given concepts while the low-scoring students used incorrect connections or did not connect concepts to the given concepts.

Research Question 5: What differences and similarities can be observed between student concept maps intended to assess structural knowledge and student responses to an open-ended assessment intended to assess content knowledge?

When comparing all the “What is science?” maps 2 through 8, the total frequencies for NOS concepts shows that the NOS concepts, Subjective and Theory, were the most frequent concepts identified in a knowledge structure (Figure 15). In addition to the NOS content, model was a concept that was highlighted as students created models and talked about how models were useful in the discovery of DNA. Models were not mentioned in the responses of the pre-instruction VNOS, but were mentioned seven times in five post-instruction VNOS. In three occasions, one student mentioned models. Regardless, the increase of the NOS concept Model observed in the “What is science?” Map 8 could be from the DNA model reference from class activities. In addition, Theory and Law concepts were not consistently used in knowledge structures throughout the 16-week semester (Figure 15). The increases and decreases between
the concepts in the knowledge structures can help describe why some students were either incorrect or confused with the following NOS tenets, inferences, tentative, myth of scientific method, per the pre-instruction and post-instruction VNOS (Table 2 and Table 3, respectively). Also, students created a new concept map each time which may suggest why concepts increase and decrease using serial mapping or that the concept is difficult for them to retain in their knowledge structure.

Shavelson et al., (2005) suggests that using concept maps provide an idea of how content in knowledge structures is represented (p. 425). The information that is provided in a map created by a student may be different than their answers on multiple-choice assessments. Burrows and Mooring (2015) recognized the importance of using appropriate assessments and implemented multiple methods to determine how chemistry students related key concepts. Using multiple assessments allowed the researchers to better understand what relationships students made that may not be represented in a concept inventory. Similarly, the VNOS was used as a previously validated open response tool. While interviews were not completed using the VNOS as its protocol requires (Abd-El-Khalick & Lederman, 2000), the open response questions allowed researchers to review what ideas were mentioned. Furthermore, Abd-El-Khalick & Lederman (2000) recognize that the NOS tenets should be well-integrated within science knowledge structures as they provided a detailed concept map of how NOS can be embedded in science concepts (Figure 1, p. 1064). Thus, using the concept maps with the VNOS open response tool can provide further evidence of what gaps may exist or how knowledge structures change throughout the semester. In addition, these assessments and tools can be used on their own in a classroom or as research, but they can complement each other by providing relevant information of concepts and connections of the concepts.
Limitations

The study was limited to one semester of data collection using multiple assessments from seventeen participants who were registered in a one semester science course. The results of this study may not apply to every science course or group of students. Furthermore, the focus of the research was the quantitative measurements from the maps students created. Thus, the quality of the connections that were made were not assessed. However, because the group was similar in context and several assessments were used throughout the semester, the variability was limited.

Recommendations for Further Study

More research is necessary when completing serial concept maps and using a variety of concept maps. For instance, monitoring students when given previously completed maps like Quinn et al., (2003-04) and Martin et al., (2000) would the knowledge structure complexity consistently increase throughout the semester? Thus, researchers may want to ask students to revisit prior maps to reflect or add upon their knowledge structure represented in a concept map. Or, similarly to Anohina-Naumeca (2015) by providing 3-5 concepts from the previous maps and determine its influence upon concepts knowledge structure complexity. Considering the many variations how concept mapping can be implemented in the classroom, the assessment tool provides many opportunities on how to monitor students’ knowledge structure. It is recommended that the study be completed in other content areas to determine how concept maps can be used to monitor students’ knowledge structure of different types of content. To complement future concept mapping activities, researchers or teachers may ask students to complete the mapping activities using a think aloud protocol to monitor how those connections are being made as similarly done in Burrows and Mooring (2015). Or, students complete a card sorting exercise that will ask them to associate different concepts together. Using multiple tools
in research and classrooms may be helpful in providing a richer assessment profile. Finally, concept mapping provides a variety of multiple assessments can be useful in creating content or thematic profiles of students if they are a novice or an expert.

Conclusions

Based on this study, the concept map training of how to create concept maps is critical to students creating maps that are reflective of their knowledge structure and useful for quantitative measurements. Using a variety of maps in the classroom can help answer questions regarding students’ structural knowledge and assessment design. Specifically, the directed maps were limited by the amount of information students provided in their maps and therefore might artificially restrain the assessment. However, while providing more direction to the task might limit student concept maps, asking more directed or specific questions might produce more detailed maps. In addition, this research suggests that it is possible to identify whether or not students have made connections across content using maps that are asking different, but related questions. Finally, after evaluating multiple maps that were collected and analyzed, the various mapping tasks do accurately reflect knowledge exhibited by students using other forms of assessment (open response). The concept maps can be used to compliment other types of assessments that do not as easily show how students identify relationships between concepts.
REFERENCES


Appendix A. “What is science?” Concept Map

“What is science?”

Compile a list of important terms that you think of when answering the question, “What is Science?”

Rank, using numbers, the concepts listed above from most general to least general.
Now, connect concepts from your list, one pair at a time, with directional links; and most importantly, to label the linking lines using short phrases. You may use cross linking, many branches and many levels of hierarchy. Continue this process until all the concepts appear on your map. Remember the boxes or ovals should contain only one or two words.
Appendix B. NOS Quiz Concept Map

Scientific Inquiry
QUIZ 1- PART A

Name_____________________________  Section: _____________________

Date: ______________________________

Directions: Please read each question carefully. This quiz is worth 10 points (2 points are from part a). You will be given 20 minutes to take it. If you finish early, please bring it to your instructor and sit QUIETLY until the rest of the class is finished.

Using the four NOS activities we learned in class, which one was the most useful in understanding aspects of NOS? Write your selection below on the line provided.

Twirlie  Fossil  Tube  Cube

__________________________________

List the concepts of nature of science that you learned from the activity chosen above.

Rank, using numbers, the concepts listed above from most general to least general.
Now, connect concepts from your list, one pair at a time, with directional links; and most importantly, to label the linking lines using short phrases. You may use cross linking, many branches and many levels of hierarchy. Continue this process until all the concepts appear on your map. Remember the boxes or ovals should contain only one or two words.
Appendix C. Mitosis (Directed) Concept Map

Make a concept map using the following terms: Mitosis, Chromosome, Chromatid, Interphase, Cell Cycle, Synthesis, Metaphase, Centromere, Spindle Fibers, Daughter Cells, Centrosome.
Appendix D. “What is biotechnology?” Concept Map

Focus Question: What is biotechnology?

Compile a list of important terms that you think of when answering the question.

Rank, using numbers, the concepts listed above from most general to least general.
Now, connect concepts from your list, one pair at a time, with directional links; and most importantly, to label the linking lines using short phrases. You may use cross linking, many branches and many levels of hierarchy. Continue this process until all the concepts appear on your map. Remember the boxes or ovals should contain only one or two words.
Appendix E. Views of Nature of Science (VNOS)

**VNOS-270**

Name: ____________________________________

Date: ________________________________

Previous science courses:

- Bios 1700  Life Science for Elementary Educators 1
- Phys 1800  Physical science for Elementary Educators 1
- Geog 1900  Earth Science for Elementary Educators 1
- Bios 2700  Life Science for Elementary Educators 2  [current]
- Sci 2800  Physical Science for Elementary Educators 2
- Geos 2900  Earth Science for Elementary Educators 2

Other college level science courses? Please list: ________________________________

Please answer each of the following questions. You can use all the space provided to answer a question. Some questions have more than one part. Please make sure you write your answers to each part in the appropriate space provided.

This assignment will not be graded. There are no “right” or “wrong” answers to the following questions. I am only interested in your ideas relating to the following questions.

1. **What, in your view, is science?** How can you determine when something is science (such as biology or physics) and when something is not science (such as religion or philosophy)?

   __________________________________________________________________________

   __________________________________________________________________________

   __________________________________________________________________________

   __________________________________________________________________________

   __________________________________________________________________________

2. **How are science and art similar? How are they different?**

   __________________________________________________________________________

   __________________________________________________________________________

   __________________________________________________________________________
3. Scientists agree that about 65 millions of years ago the dinosaurs became extinct. However, scientists still disagree about what caused this extinction. Why do you think they disagree even though they all have the same information?

__________________________________________________________________________________
__________________________________________________________________________________
__________________________________________________________________________________
__________________________________________________________________________________

(b) Do you think this controversy could be resolved? If so, how? If not, why not?

__________________________________________________________________________________
__________________________________________________________________________________

(c) How do you think scientists know how dinosaurs looked and moved?

__________________________________________________________________________________
__________________________________________________________________________________

4. There are many types of phenomena (past, present, and future) that scientists study, but cannot see. For example, scientists have never seen “dark matter”, the center of the earth, or into the nucleus of an atom. Yet many scientists use their understanding of these phenomena to do research.

If they have never seen these things, what kind of information do scientists use to figure out these things exist or what they look like?

__________________________________________________________________________________
__________________________________________________________________________________
__________________________________________________________________________________

(b) Should we, as a public, accept scientists’ explanations or descriptions of things they have not seen? Why or why not?

__________________________________________________________________________________
5. Scientists try to find answers to their questions by doing investigations. Do you think that scientists use their imagination & creativity in their investigations?

(a) If you think "YES", explain why and in what part of their investigations (planning, analysis of data, interpretation, etc.) you think they use their imagination & creativity.

(b) If you think "NO", explain why imagination & creativity are not part of science.

6. What do you think is the difference between a scientific theory and a scientific law?

A scientific theory is....

A scientific law is....

(b) Give an example of a scientific theory and an example of a scientific law.

Example of a Scientific Theory:

Example of a Scientific Law:
(c) Do you think scientific theories we have today will change in the future? Why or why not?

___________________________________________________________________________________________

___________________________________________________________________________________________

(d) Do you think scientific laws we have today will change in the future? Why or why not?

___________________________________________________________________________________________

___________________________________________________________________________________________
Date: March 24, 2017

To: Brandy Pleasants, Principal Investigator
Leah Cook, Student Investigator for dissertation

From: Amy Naugle, Ph.D., Chair

Re: HSIRB Project Number 17-03-21

This letter will serve as confirmation that your research project titled “Assessing Structural Knowledge in a Science Classroom” has been approved under the exempt category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note: This research may only be conducted exactly in the form it was approved. You must seek specific board approval for any changes in this project (e.g., you must request a post approval change to enroll subjects beyond the number stated in your application under “Number of subjects you want to complete the study”). Failure to obtain approval for changes will result in a protocol deviation. In addition, if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

Reapproval of the project is required if it extends beyond the termination date stated below.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: March 23, 2018