THE DEVELOPMENT AND USE OF A CONCEPT MAPPING ASSESSMENT TOOL WITH YOUNG CHILDREN ON FAMILY VISITS TO A LIVE BUTTERFLY EXHIBIT

By

JENNIFER CHERYL MESA

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2010

To my family and friends

ACKNOWLEDGMENTS

I thank my ever-patient and supportive husband, John Mesa, and dear friends, Michelle Klosterman, Katie Milton, and Mary Perkins. I also thank my chair, Linda Cronin-Jones, and committee members, Tom Dana, Betty Dunckel, Dave Miller, and Barbara Pace, for their valuable feedback. This dissertation would have never been completed without all of your help and encouragement.

TABLE OF CONTENTS

page

CHAPTER

APPENDIX

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

THE DEVELOPMENT AND USE OF A CONCEPT MAPPING ASSESSMENT TOOL WITH YOUNG CHILDREN ON FAMILY VISITS TO A LIVE BUTTERFLY EXHIBIT

By

Jennifer Cheryl Mesa

August 2010

Chair: Linda Cronin-Jones
Major: Curriculum and Instruction

Although young children are major audiences of science museums, limited evidence exists documenting changes in children's knowledge in these settings due in part to the limited number of valid and reliable assessment tools available for use with this population. The purposes of this study were to develop and validate a concept mapping assessment tool and to use this tool to document the butterfly-related knowledge of young children on unguided family visits to a live butterfly exhibit at a natural history museum. In this study, forty-two children visited the live butterfly exhibit with their families on unguided tours and completed pre- and post-visit concept mapping tasks. During pre- and post-visit mapping sessions, children created and revised concept maps about butterflies using a set of eight butterfly-related concept pictures and provided verbal explanations for each picture pair in their concept maps.

Three raters used three different scoring systems designed for use in this study to evaluate the scientific accuracy of the children's pre- and post-visit maps, including the picture pairs and verbal explanations. Quantitative analyses of the scores indicate that the raters used the three scoring systems with a moderate to high level of consistency. The results also indicate that children significantly increased their butterfly-related knowledge in the live butterfly exhibit regardless of recent prior experience with the exhibit.

Qualitative analyses of children’s verbal explanations indicate that children possessed butterfly-related knowledge related to: the needs of butterflies, the life cycle of butterflies, the ecology of butterflies and other insects, the diversity and classification of butterflies and other insects, the threats to and conservation of butterflies, and the social and maternal behavior of butterflies. Although children with different levels of exhibit experience showed similarly high levels of prior and subsequent knowledge, the types of understandings they communicated in their verbal explanations differed somewhat. Surprisingly, children without recent prior exhibit experience showed greater understanding in more areas of butterfly-related knowledge than children with recent prior exhibit experience. The results of this study have implications for the field of science education in general and the field of informal science education in particular.

CHAPTER 1
INTRODUCTION

Children do not begin learning about science when they enter the doors of a school. They first learn about science in their homes and communities with their families. At home, young children learn science as they participate in everyday experiences such as interacting with media (e.g., books, television, and internet resources), exploring various materials through play behavior, and making observations of the natural world. In the community, young children learn science as they accompany their parents on errands and take family trips to informal settings such as nature parks, science museums, and zoos. Even the commonplace visit to the grocery store can be a science learning experience as young children learn the rudiments of nutrition and practice their observation skills in the produce aisle.

For many young children, family visits to local science museums or zoos are commonplace events in their lives. The increasing popularity of science museums as sites of choice for family recreation and learning is affirmed by the surge in construction of new children's museums and continued increases in science museum attendance rates in the U.S. (Association of Children's Museums, 2009). Children's museums commonly include exhibits that focus on both science content and process skills. In 1975, there were approximately 38 children's museums in America while 243 exist today. Furthermore, an additional 78 children's museums are currently in the planning stage throughout the country.

Similarly, museum attendance in the U.S. has increased to the point where it is estimated that one in five Americans visited a science museum in 2008 (Association of Science and Technology Centers, 2009). Families account for more than half of science museum visitors and many science museums have developed special exhibits and programming for young children. Such exhibits and programming are costly investments for both science museums and families on tight budgets in these economic times.

Research indicates that families visit science museums with the expectation that such experiences are educational for their young children (Ellenbogen, 2002; Falk & Dierking, 1992; Hilke, 1989; Hood, 2004). Given that the education of young children is part of the stated mission of many science museums, these informal settings do little to dissuade families of this notion. Science museums also obtain targeted funding from many governmental and non-governmental organizations that is specifically designed to maintain and improve their offerings for young visitors. For example, the Children's Museum of Boston recently obtained a $1.5 million grant to develop a new science and math exhibit for three-to-five year old children and their families from the Division of Research on Formal and Informal Settings of the National Science Foundation (National Science Foundation, 2007).

Despite the widespread agreement that science museum visits should be educational for visitors of all ages, few methods or tools for documenting the science learning of young children in such informal settings currently exist. The development of additional learning assessment methods and tools may allow science museums to more effectively advance their understanding of the nature of science learning occurring among young visitors and promote testing and evaluation of new ways of enhancing the learning experiences of young museum visitors.

Purpose

Despite the current popularity of science museums in the U.S., little is known about what young children learn on family visits to these informal settings. In this new age of educational accountability, science museums and other informal science learning institutions cannot afford to assume that their visitors are actually learning. Instead, science museums need to clearly document the measurable learning outcomes of all of their visitors, including young children, in order to obtain and maintain funding sources. Given this reality, any claims of learning in science museums must be substantiated using valid and reliable assessment tools. Thus, this study was designed to address this need by attempting to document the learning of young children on family visits to a live butterfly exhibit at a natural history museum. In order to accomplish this goal, a new concept mapping tool for assessing the science learning of young children (aged five to seven years) in informal settings was developed and tested as part of this study. Specifically, the following questions were addressed in this study:

1. Which of three scoring systems can be used the most reliably by raters to evaluate the concept maps of young visitors to a live butterfly exhibit?
2. What is the validity of using concept maps to assess the butterfly-related knowledge of young visitors to a live butterfly exhibit? To answer this question, the following two sub-questions were posed. These questions focus on the feasibility and appropriateness of using concept maps to assess young children's butterfly-related knowledge in a live-exhibit informal setting:
    a) To what extent can young children construct and verbally explain propositions in their concept maps?
    b) To what extent do young children interpret the concepts in their maps as intended?
3. How does young children's butterfly-related knowledge change on family visits to a live butterfly exhibit as measured by the concept mapping assessment tool?
4. How do the concept maps of young children with and without recent experience with a live butterfly exhibit compare in terms of accuracy and content as measured by the concept mapping assessment tool?

Statement of Need

Although families are consistently a major audience of science museums, few studies have focused on the learning of children during family museum visits. The learning of young children has been particularly understudied, perhaps because of the numerous challenges associated with studying this population. The challenges of conducting research with young children in museums are twofold: first, these children often lack the language skills needed to understand the questions posed by researchers, and second, the museum environment is often not an ideal research setting. The effects of confounding variables are particularly difficult to isolate and control for in museum research as the experiences of visitors vary with their learning behavior decisions they make while visiting exhibits and are out of the control of researchers (Allen, 2008). Despite these major challenges, the need for studies investigating the learning of young children in science museums is also twofold. First, positive museum learning experiences during early childhood may stimulate lifelong interest in science and foster lifelong science learning; and second, science museums are increasingly investing money and resources to develop experiences and exhibits tailored to the specific needs and interests of families with young children.

Anecdotal evidence suggests that early visits to informal science learning institutions have great potential to awaken and sustain long-term interest in science. Many scientists and engineers have acknowledged the important influence of early visits to science museums and zoos on their career choices (Cosmos Corporation, 1998). Furthermore, Miller (2004) reported that adults engaged in Science-Technology-Engineering-Mathematics (STEM) careers participate in informal science learning experiences such as science museum visits more frequently than adults engaged in non-STEM careers. These individuals likely visit science museums with their spouses and children. This is significant because some evidence suggests that adults who visited science museums as children are more inclined to take their own children to museums and view museum visits as a worthwhile use of their leisure time (Hood, 2004). A clearer understanding the role science museum visits play in the lifelong science learning of individuals is needed, especially during early childhood when science museum visits are typically family events. In order to document the impacts of early family visits to science museums on short-term and lifelong science learning, additional valid tools for measuring the science learning of young children must be developed.

Other researchers have suggested that informal science learning experiences foster the development of interest and expertise in science during early childhood (Crowley & Jacobs, 2002; Palmquist & Crowley, 2007). Such experiences not only include visits to informal science learning institutions but also other experiences in the home and community such as watching television, reading books, and going to the park. Researchers are beginning to investigate the role of these experiences in young children's understanding of science. In the past, many science education researchers have assumed that family museum visits contribute to the scientific understanding of children and adults on such visits; however, to date, older children (older than eight years old) and adults have been the main focus of most research. Clearly, additional studies focusing on the learning of young children on family museum visits are needed.

Meanwhile, educational programming targeting families and young children has become increasingly popular in science museums. Many science museums have invested significant resources developing exhibits, learning playgrounds, discovery rooms, and workshops specifically targeting families and young children. However, the published literature documenting how these exhibits and programs affect learning is limited (Borun, Chambers, & Cleghorn, 1996; Borun, Chambers, Dritsas, & Johnson, 1997; Borun & Dritsas, 1997). The lack of published research in this area may be due in part to the difficulties associated with assessing the learning of young children in museum settings.

The few studies that have investigated the learning of families with young children at science museums have examined how parents and young children interact with specific exhibit features and with each other. These studies have also highlighted how parents scaffold the science understanding of their children during science museum visits (Ash, 2004; Blud, 1990a; Crowley, Callanan, Jipson, et al., 2001; Crowley, Callanan, Tenenbaum, & Allen, 2001). Through observations of family learning behaviors and conversations at exhibits, researchers have found that young children do learn more on science museum visits when parental scaffolding occurs.

Although observational studies are useful tools for determining what parents and children do and say when interacting with science museum exhibits, observational studies alone may not provide all of the data needed to fully determine how and what young children learn from museum visits. Learning, after all, is an internal process of constructing meaning from one’s experiences over time. Learning in science museums is influenced by many factors, including visitors’ personal characteristics (e.g., age, personal motivations, and prior knowledge), social interactions at exhibits, and the physical environment of the museum (Falk & Dierking, 1992). Observational studies can only provide a partial picture of what and how young children learn during science museum exhibits.

Likewise, studies focused on analyzing the spontaneous verbalizations, questions, and other aspects of conversations between young children and their parents during museum visits can only provide a partial picture of the museum learning process. Clearly, in order to more fully understand and document the learning of young children during science museum visits, additional sources of evidence examining young children's cognitive structures are needed. These additional data sources could provide a more complete picture of young children's learning and provide stronger support for researchers' inferences and conclusions.

The narrow range of research methodologies employed by researchers investigating museum-based learning was recognized by the National Association for Research in Science Teaching's (NARST's) "Informal Science Education" Ad Hoc Committee when it issued a call to explore alternative methodologies for investigating the learning of visitors to museums (Dierking, Falk, Rennie, Anderson, & Ellenbogen, 2003). To address the clear need for development and evaluation of new research methodologies and tools for use in science museum contexts, this study developed and tested a new concept mapping tool designed to assess the museum learning of young children.

In formal educational settings, concept mapping has been used as an assessment tool in preschool and primary grade classrooms with some success (Figueiredo, Lopes, Firmino, & de Sousa, 2004; Fleer, 1996; Hunter, Monroe-Ossi, & Fountain, 2008). The validity and reliability of these applications has not been well documented, however. In science museum settings, concept mapping has shown promise as a tool for documenting older children’s thinking about science content (Anderson et al., 2000). Again, the reliability and validity of the concept mapping assessment tools designed for this purpose have not been well documented. Thus, this study evaluated the reliability and validity of a concept map assessment tool designed specifically for use with young children in a museum setting.

Significance of the Study

Currently, there are limited techniques and tools available for examining and documenting the learning of young children in science museum settings. This study explored the use of a new tool for studying the learning of young children in science museums and thus may provide researchers and museum educators with a new, effective method for studying how family visits to science museums impact the cognitive knowledge structures of young children. Similarly, using a concept map assessment tool with young children may allow researchers to more thoroughly study the role of various aspects of family science museum visits such as interactions between family members and exhibit features. A more thorough understanding of how various aspects of a family museum visit impact the science learning of young children may help museum professionals develop more educationally-effective programming. Developing techniques for enhancing the museum-based science learning of young children has implications for their future science achievement in school and may even influence their pursuit of Science-Technology-Engineering-Mathematics (STEM) careers.

Other museum researchers may be able to use the findings of this study to develop or modify concept map assessment tools for use with other difficult-to-study populations, such as non-English speakers or learners with particular verbal limitations or disabilities. Considering the increased diversity of languages spoken in many parts of the United States, some museums are now serving populations of visitors who cannot speak English, let alone read or write it. The concept map assessment tool developed for this study may be able to serve as a valid, reliable, language neutral tool for studying these populations.

Definitions

For the purpose of this study, the definitions below are provided to illuminate key terms.

Concept: A perceived regularity in objects or events in the natural world which is identified by words or other symbols (Novak & Gowin, 1984; Novak & Canas, 2007).

Concept map: A graphic tool used to organize and represent knowledge, particularly the interrelationships between concepts (Novak & Gowin, 1984; Novak & Canas, 2007). Concept maps are comprised of propositions often organized in a hierarchical fashion.

Informal science learning institution: A place that is specifically designed to support the science learning of visitors (Bell, Lewenstein, Shouse, & Feder, 2009). Examples include science museums, zoos, aquaria, and nature centers.

Proposition: A statement about the relationship between two concepts (Novak & Gowin, 1984; Novak & Canas, 2007). Traditionally, a proposition is formed when two concepts represented by words or phrases are connected by a line (or arrow). Linking words or phrases are placed on the line (or arrow) to describe the relationship between the two concepts.

Summary

Young children with family groups are a significant part of the target audience for many science museums. As a result, many of these institutions have developed programming and exhibits specifically targeting this population. However, the body of existing research literature includes few studies documenting the impact of science museum programs and exhibits on the learning of young children. Furthermore, early childhood has been recognized as a key period in the development of science-related interest and knowledge. For this reason, the impact of early experiences in science museums on the development of early interest and expertise in science should be investigated more thoroughly.

Researching the learning of young children in science museums poses significant challenges due to the limited language abilities of children and the nature of these informal settings. In the past, researchers have relied primarily on observations and interviews as data sources when studying learning in science museums. Although these descriptive data sources have provided useful information about the interactions of young children and families at science museum exhibits, they have not sufficiently described and documented exactly what science content young children learn during these visits or how this content is cognitively organized by young children. Alternative methods of assessing the learning of young children at science museums are needed. This study aimed to develop and evaluate a new museum research tool while also contributing to the currently sparse body of literature regarding the learning of young children during family visits to a live butterfly museum exhibit.

CHAPTER 2
REVIEW OF LITERATURE

Theoretical Framework

The Contextual Model of Learning conceived by Falk and Dierking (1992) was used as a framework for examining the learning of the young children participating in this study. This model is especially useful when studying learning in informal settings such as science museums and is based on observations of museum visitors and guided by research in psychology, neuroscience, and anthropology. In the view of Falk and Dierking (2000), learning is “an adaptation enabling people to intelligently navigate an ever-changing social, cultural, and physical world” (p.13). Accordingly, the Contextual Model of Learning is aligned with the social constructivist perspective and describes learning as an active process in which people continuously strive to make meaning from social experiences in a variety of settings throughout their lives. In the model, the constant interplay of three contexts, the personal, the socio-cultural, and the physical, influences learning. These contexts are not mutually exclusive as some elements of each may be represented in more than one context. The following discussion highlights aspects of these three contexts particularly relevant to this study.

The Personal Context of Learning

The personal context of learning is composed of the diverse motivational and emotional cues that influence learning (Falk & Dierking, 1992; 2000). Within the informal learning environments of museums, individuals are generally intrinsically motivated to learn. They usually attend museums by choice and they receive no rewards other than the pleasure of the experience. Furthermore, interest, the psychological construct that includes attention, persistence in a task, and continued curiosity, acts as a filter for the abundance of sensory input at museums (Falk & Dierking, 2000). Visitors pay attention to what interests them, and personal interests frequently arise from positive prior experiences with topics. The positive feelings, attitudes, and emotions surrounding prior experiences promote visitor interest, which in turn influences the specific museums and exhibits they choose to visit.

Similarly, the understandings gained from previous experiences give visitors the ability to make connections at exhibits and experience success in learning, which further promotes interest. This interest is demonstrated by their decisions to visit specific museums or pay selective attention to specific exhibits. These choices allow visitors to maximize their learning experiences in museums by using these experiences to reinforce and build on what they already know. A positive feedback loop emerges in which visitors use museums to support deepening interest in and knowledge of a domain area. In this way, museum-going can become cemented as a learning experience in the minds of museum enthusiasts (Ellenbogen, 2002). One aspect of this study explored the impact of recent prior experiences with a live butterfly exhibit on the understanding of young children on family visits to the exhibit.

The Socio-cultural Context of Learning

The socio-cultural context of learning includes the wide variety of social events that mediate learning (Falk & Dierking, 1992; 2000). In museums, visitors often interact with each other and with socially-constructed tools, signs, and symbol systems. Each visitor brings to these interactions a unique complement of values, beliefs, and norms that influence his/her perceptions and behavior. These values, beliefs, and norms are cultural products that are transmitted across generations from parent to child. For example, when parents take their children to museums to “do the museum,” experience a new exhibit, or use the reading room; they are demonstrating the value of museums as places for learning. In addition, parents model the norms of being a museum visitor, including ways of interacting with exhibit features, docents, and other visitors, for their children.

Parents also play a large role in mediating the development of social and cognitive skills by the simple act of talking to their children. According to Falk and Dierking (2000), “this social interaction happens constantly, during mealtimes, during visits to museums, and in other free-choice learning settings, even while supposedly reading the newspaper alone” (p.45). Such conversations provide opportunities for children to revisit and make sense of their experiences, particularly those that are part of the family history. Furthermore, parents have been found to scaffold their children’s investigations in museums in many ways, including questioning, activating prior knowledge, and guiding problem-solving (Ash, 2003; Crowley, Callanan, Jipson, et al., 2001; Ellenbogen, 2002).

This social support allows children’s learning to take place within a “zone of proximal development” (Vygotsky, 1978). Parents’ skill in scaffolding their children’s learning enables children to experience higher levels of learning that they would not be able to experience alone. For this reason, this study examined the learning of children in the company of their parents rather than studying children in school groups. Although children in school groups may also benefit from the scaffolding of a teacher or chaperone, the conversations between children and parents are more likely to include references to shared history. These references are important for children’s meaning-making by activating prior knowledge (Ash, 2003, Crowley, Callanan, Jipson, et al., 2001; Ellenbogen, 2002).

The Physical Context of Learning

The physical context of learning consists of the physical environment in which learning takes place. Museums are designed places where people specifically come to “see real objects, placed with appropriate environments” (Falk & Dierking, 2000, p.139). Museum designers manipulate many features of the physical environment to “visually, and increasingly aurally and socially, attract, and pull in the visitor” (Falk & Dierking, 2000, p. 123). However, the plethora of sights and sounds in museums can lead to sensory overload in some visitors, especially those who are unfamiliar with museum settings. Sensory overload quickly reduces the ability of inexperienced visitors to process the information in exhibits.

Alternatively, visitors with more experience in museum-going or the subject matter at hand have a greater ability to process and “chunk” the information presented in exhibit displays (Falk & Dierking, 2000). This means that they are able to group items in exhibit displays in meaningful ways that can easily be remembered. Inexperienced visitors often just see a hodgepodge of unrelated objects and text in an exhibit and can recall very little of what was present. Design features such as clustering conceptually-related exhibits and explicit labeling of exhibit messages can make the content more comprehensible for inexperienced visitors (Falk, 1997). Although the designers of the live butterfly exhibit that was the focus of this study used explicit labels on their signage, the children in this study had limited access to the information on these labels unless their parents read them aloud. In addition, the live butterfly exhibit contains an impressive array of sensory experiences which may contribute to sensory overload, which provided another reason for exploring the impact of recent prior experiences on the learning of children participating in this.

Research Framework

Mixed methods research is “the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts, or language into a single study” (Johnson & Onwuegbuzie, 2004). Many leading researchers in the field of informal science education increasingly advocate the use of mixed methods research which is thought to have the potential to capture the full range of visitor learning and take into account the personal, socio-cultural, and physical contexts of learning (Allen, 2008; Dierking et al., 2003; Rennie & Johnston, 2004). Furthermore, mixed methods research approaches may be better suited for the challenges of conducting research in informal science education settings by allowing researchers to investigate a broader range of research questions than can be investigated when solely quantitative or qualitative approaches are used.

In museum research, challenges arise from the great diversity of visitor learning experiences that can occur at exhibits. Visitors make innumerable choices about how they interact with museum exhibits and with each other. These choices in turn affect the quality of their learning experiences. Purely quantitative research may restrict the types of questions that can be investigated regarding visitor learning experiences while purely qualitative research may rely too heavily on inferences when interpreting visitor behaviors and learning (Rennie & Johnston, 2004).

When studying learning in informal settings, Rennie and Johnston (2004) support methods that allow researchers to “see through the eyes of visitors” (p. 8). In other words, they contend that the best methods for conducting research in museums are those that allow the researcher to directly speak to visitors about their experiences. Aligned with these views, this study employed a mixed methods design in which I, as the researcher, directly interacted with the child participants in the museum setting.

Characteristics of Learning in Museums

In the last decade, the informal science education community has come to the consensus that all learning is personal, depends on the context, and takes time (Falk & Dierking, 1992; 2000; Rennie & Johnston, 2004). Informal science education researchers further assert that the distinction often made between learning that occurs in classrooms and informal environments such as museums is artificial. More specifically, they reason that learning “occurs across a broad spatial and temporal context, both inside and outside schooling” (Dierking et al., 2003, p.110).

Learning is a cumulative process that involves many seemingly minor changes in knowledge, attitudes, and behavior arising from various experiences in and out of the traditional classroom. According to Dierking et al. (2003), “most learning has more to do with consolidation and reinforcement of previously understood ideas than the creation of totally new knowledge structures” (p.110). Considering this view, museums and other informal science learning environments have much to contribute by both introducing and reinforcing ideas about science that are essential for learning science in the classroom.

Museums have historically been the focus of much of the research on science learning outside of school. Many researchers have recognized that the museum context is a unique context for learning about science. First, researchers have noted that the museum context offers visitors innumerable choices, ranging from the choice of whether or not to visit a museum to the choice of whether to engage in a particular learning opportunity in a museum (Rennie, 2007; Rennie, Feher, Dierking, & Falk, 2003; Rennie

& Johnston, 2004). This differs significantly from the traditional classroom context where learners are obligated to come to class and have little to no choice in the selection of learning experiences and activities.

Second, free-choice learning opportunities in the museum context are non-evaluative and non-competitive, unlike many learning opportunities in the traditional classroom context (Rennie, 2007). There are usually no extrinsic rewards or consequences for participating in learning opportunities at museums. Third, the social interaction that occurs in the museum context is also significantly different from the social interaction that takes place in the classroom context (Rennie, 2007). Family groups comprise the main audiences of many museums and include members of varying ages and abilities. It is not unusual to find three generations of a family present at a museum exhibit. The children and adults in a family each have a role in determining the “family agenda” of a museum visit (Ellenbogen, Luke, & Dierking, 2004). In contrast, the social interaction in traditional classroom contexts is “constrained between same-age peers and formalized with the teacher as the main adult” (Rennie, 2007, p.127). Thus, the teacher assumes primary responsibility for setting the agenda for learning in the traditional classroom context.

Learning in Museums Research

When the body of research on learning in museums is viewed as a whole, it is clear that great progress has been made in determining the influences of personal, socio-cultural, and physical contexts on learning in museums (Falk & Dierking, 1992; 2000). With regard to the personal context, the influences of visitor characteristics as well as prior knowledge and interest in a topic have been explored and documented by several researchers (Anderson & Lucas, 1997; Anderson et al., 2000; Boisvert & Slez, 1994; Falk, Moussouri, & Coulson, 1998; Greenfield, 1995; Kubota & Olstad, 1991). Other researchers have explored and compared solitary and shared learning in museums and have focused on the socio-cultural context of learning (Ash, 2003; Cox-Petersen, Marsh, Kisiel, & Melber, 2003; Cox-Petersen & Pfaffinger, 1998; Crowley, Callanan, Jipson, et al., 2001; Gilbert & Priest, 1997; Packer & Ballantyne, 2005; Sandifer, 1997; Tunnicliffe, 2000; Tunnicliffe, Lucas, & Osborne, 1997). Finally, still other researchers have documented how exhibit design, an aspect of the physical context, influences visitor learning in museums (Allen, 1997, 2004; Borun et al., 1996; Borun, Chambers, et al., 1997; Borun & Dritsas, 1997; Chermayeff, Blandford, & Losos, 2001; Falk, 1997; Schauble & Bartlett, 1997).

The Personal Context of Learning in Museums

Each visitor entering a museum differs in his/her personal characteristics and prior experiences. Personal characteristics influencing learning include age and gender as well as motivations for visiting a museum. Boizvert and Slez (1994) documented the behavior of a purposeful sample of 154 visitors at a human body discovery room at a science museum and observed that older adolescents (ages 15-18) and young adults (ages 19-24) spent less time viewing exhibits compared to other visitors. Similarly, young adults demonstrated lower levels of engagement with the exhibits. Boizvert and Slez found no differences in the amount of time male and female visitors spent at exhibits or their levels of engagement with exhibits. However, Greenfield (1995) observed over 2000 visitors at a problem-solving based life and physical science exhibition at a natural and cultural history museum and found evidence that the type of exhibit is related to the proportion of males and females interacting with an exhibit.

Specifically, Greenfield found that boys used physical science exhibits to a greater extent than girls.

Visitor motivations

With regard to motivations, visitors attend museums for different reasons which impact what they do and learn on their visits. Falk et al. (1998) used interviews, questionnaires, and a modified concept mapping technique called personal meaning mapping to determine how the personal agendas, that is, the motivations and visit strategies, of 40 adult visitors affected their learning at a geology exhibition at a natural history museum. Falk et al. documented six different categories of motivations representing the full range of perceptions people have about museums and their potential for leisure and learning. Interestingly, they noted that most visitors have more than one motivation for visiting museums. In general, individuals with strong educational motivations show significant conceptual learning while individuals with strong entertainment motivations show significant vocabulary development. Individuals with both strong educational and strong entertainment motivations show significant vocabulary and conceptual development. Furthermore, visitors with strong educational motivations spend longer periods of time in exhibits compared to visitors with weak educational motivations.

Visitor prior experiences and knowledge

The prior experiences of visitors may promote different levels of exploratory behavior and learning at exhibits. Several researchers have explored the impact of novelty of museum visits on the exploratory behavior and learning of schoolchildren (Anderson & Lucas, 1997; Anderson et al., 2000; Kubota & Olstad, 1991). High levels of novelty are thought to hinder learning at exhibits by encouraging diversive exploration rather than specific exploration (Kubota & Olstad, 1991). Diversive exploration constitutes exploration of the entire museum environment instead of focused exploration of specific exhibits or exhibit features. These studies have attempted to reduce novelty by providing children with vicarious experiences of museum exhibits in the form of slide show or video presentations. Researchers content that these pre-visit orientations can reduce novelty and improve children's prior knowledge of the physical environment of a museum and appropriate strategies for interacting with the exhibits.

Kubota and Olstad (1991) compared exploratory behavior and post-visit knowledge of 64 sixth-grade children from intact classes at one school who visited a science playground at a science center. Before visiting the science playground, half of the children were randomly assigned to view a slide show of similarly aged children interacting with features of the science playground while the other half in the control group viewed a slide show of an unrelated exhibit. Boys who received the relevant pre-visit orientation demonstrated more specific exploratory behavior at the exhibit and gained greater knowledge of exhibit content on a paper-and-pencil post-test than boys who received the non-relevant slide show. However, in this study, participating in a relevant pre-visit orientation appeared to have no effect on the exploratory behavior or knowledge gains of girls. As the authors pointed out, the content focus of the science playground was related to physical science, which may have influenced the lower levels of exploratory behavior and learning gains of the girls.

On the other hand, Anderson and Lucas (1997) found no gender differences in their study of 75 eighth-grade children's knowledge of physical science exhibits at a science center. Similar to the Kubota and Olstad (1991) study, the children in the

Anderson and Lucas study were from intact classes at one school and half of the children were randomly assigned to watch a video about the science center’s exhibits while the other half in the control group watched an unrelated video. The children’s knowledge and relevant prior experiences were documented in paper-and-pencil post test.

To build on this prior work, Anderson et al. (2000) used interviews and concept mapping to examine how 12 eighth-grade students construct knowledge of electricity and magnetism resulting from experiences in the same science center and related in- and out-of-classroom experiences both before and after science center visit. They concluded that careful planning of pre-visit, during visit, and post-visit activities is essential for students to explore museum exhibits effectively and construct accurate understandings of science content. Although the approach of this study supports a greater understanding of how school and museum experiences relate to each other, these results have limited generalizability due to the study’s small sample size.

Falk and Adelman (2003) studied 100 adult visitors on casual trips to an aquarium. They used interviews and personal meaning mapping to compare how various levels of prior knowledge and interest impacted visitor’s subsequent understanding of exhibit content. Their results indicated that visitors with different levels of prior knowledge and interest had differing gains in knowledge. However, Falk and Adelman also noted that regardless of the extent of prior knowledge, only visitors possessing moderate to high levels of interest showed significant gains in knowledge of exhibit content. Falk and Adelman’s findings suggest the need to assess visitor learning by subgroups based on levels of prior experience rather than focusing exclusively on demographic subgroups.

Much of the existing research examining the personal context of learning in museums has been conducted with schoolchildren and adults. In general, the findings suggest that personal characteristics such as age, gender, motivation, and prior experiences all have some role to play in influencing the learning of visitors at exhibits. However, it is unknown how motivation and prior experiences influence the learning of young children on school visits or family visits.

The Social Context of Learning in Museums

Perform an informal survey of visitors at any museum exhibit over the course of a week and you will quickly notice the many different types of social groups present. The types of groups present differ based on the day of the week, with school groups more common during the week and family groups more common during the weekend (Boisvert & Slez, 1994; Sandifer, 1997). The existing literature includes conflicting findings regarding the behavior of different types of social groups visiting museums. Existing research regarding the social context of museum learning is summarized in the following sections.

Time spent at exhibits

Boisvert and Slez (1994) found no difference in the amount of time solo visitors, visitors in peer groups, and visitors in families spent at exhibits. This finding is important to note because time spent at exhibits is a frequent indicator of learning in museums (Falk et al., 1998). Packer and Ballantyne (2005) also found no difference in the amount of time solo visitors and visitors with companions spent at exhibits. In contrast, both McManus (1987) and Sandifer (1997) observed that family visitors spent more time at exhibits compared to non-family visitors. However, these two studies did not describe how family visitors spent this extra time at exhibits, raising the question of whether they were actively engaged in learning behaviors during all of the time they spent in front of particular exhibits.

Conversations at exhibits

Studies comparing the conversations of family groups and school groups at museum exhibits have revealed few differences (Tunnicliffe, 2000; Tunnicliffe et al., 1997). Tunnicliffe et al. analyzed 1193 conversations between elementary children and adults at a zoo and a natural history museum. They found that the conversations of family groups and school groups proceeded in a similar fashion. For example, visitors in family groups and school groups located an animal at the zoo, named it, described its physical features and behavior, and made interpretative comments (Tunnicliffe et al., 1997). Conversations were similarly focused on the physical features of the specimens at the natural history museum. One of the few differences reported in this study was the higher number of affective comments made by school groups at both the zoo and museum. Subsequent work at a robotic dinosaur exhibit at the same natural history museum revealed similar findings (Tunnicliffe, 2000), supporting the conclusion that school groups and family groups are more similar than they are different regarding the focus of their conversations at museum exhibits.

Although Tunnicliffe explains these similarities as products of equally social agendas of school groups and family groups, she does not present data to support this claim. It may be that the presence of children in both groups is the most important factor influencing the focus of conversations. Children are natural scientists and attempt to make sense of their world by making whatever observations and inferences they can (Lind, 2005). It is not clear from these existing studies how adults support the learning of children in school and family groups when visiting informal learning settings.

Group interactions at exhibits

Other studies have described the types of interactions occurring among members of school groups and family groups that can impact children’s learning in informal (Cox-Petersen et al., 2003; Cox-Petersen & Pfaffinger, 1998; Gilbert & Priest, 1997; Lucas, 2000). Lucas (2000) conducted a naturalistic inquiry of how one teacher integrated a museum visit into the curriculum for her seventh-grade class at an all boys’ school. He explained that on the museum visit the teacher “spent a large proportion of her time encouraging boys to focus their attention on the exhibit labels, helping them to understand how particular exhibits ‘worked’ or relating the exhibit to real life application, and encouraging them to explain the purpose of an exhibit to her or their classmates” (p. 533). On a larger scale, Cox-Peterson & Pfaffinger (1998) identified four different roles that 12 teachers adopted during school visits to a natural history museum: explainer, initiator, manager, and observer. They noted that the initiator role was the most effective in encouraging children to explore more exhibit features for longer periods of time during visits, yet they reported that only half of the 12 teachers in the study displayed this role. Along a similar line of inquiry, Cox-Peterson et al. (2003) observed that docents leading 30 different elementary and middle grades school groups on museum tours rarely acted as facilitators of student learning on school tours. Instead, the docents offered walking lectures of exhibits interspersed with factual questions.

In contrast, Gilbert and Priest (1997) toured an exhibit about bread at a science museum with a class of fourth grade students and their teacher. They provided many examples of how the docent and teacher facilitated children’s learning by inviting and extending conversations with students regarding their exploration of objects in the exhibit. Parents have been observed to take part in conversations with their children in similar ways (Ash, 2003; Crowley, Callanan, Jipson, et al., 2001). For instance, Crowley and colleagues analyzed conversations between 91 parent and child pairs at a zoetrope exhibit at a science museum and found that parents encouraged their children to engage in longer, more focused explorations of the zoetrope (Crowley, Callanan, Jipson, et al., 2001). In many cases, parents helped children identify and interpret evidence, make comparisons, and generate explanations.

Research examining social interactions at museum exhibits includes adults and children and school and family groups. However, this line of research relies heavily on inferences of learning rather than on more direct measurement of learning. None of the social interaction studies conducted to date have been able to make a compelling argument that adult behavior directly influences child learning in either school groups or family groups.

The Physical Context of Learning in Museums

Museum professionals including designers, educators, and evaluators strive to develop exhibits that “do more than entertain” (Allen, 2004). In a reflection piece, Allen provided her perspective on recent design and evaluation research at her science museum. She discussed many examples of how exhibits were designed and redesigned to facilitate science learning and the types of learning outcomes that were studied. Allen identified four characteristics of successful educational exhibits: immediate apprehendability, physical interactivity, conceptual coherence, and broad appeal. These four characteristics mainly stem from a user-centered design approach and efforts to make the conceptual foundations and applications of exhibit content explicit to visitors. In her reflection, it is clear that Allen views exhibit design as an iterative process that should be responsive to the museum audience at all times.

Designing exhibits for inquiry

In contrast to Allen's (2004) overview of one museum's ongoing design and evaluation process, several other studies report on specific design initiatives (Allen, 1997; Borun et al., 1996; Borun, Chambers, et al., 1997; Borun & Dritsas, 1997; Chermayeff et al., 2001; Falk, 1997). Many of these studies demonstrate the efficacy of encouraging visitor inquiry in exhibits; however, the strength of the evidence for their claims varies. In an earlier paper, Allen (1997) described how she used interviews and two performance tasks to determine which of seven inquiry activities was the most effective in facilitating visitor understanding at an exhibit on colored shadows. Participants included 392 adults and children who volunteered to take part in the interviews while on casual visits to a science museum. The most promising activity for the adults, especially those with a college education, was an interpretation task in which they generated an explanation after interacting with the exhibit and reading a scientific explanation of the phenomena at the exhibit. In other words, the activity mirrored the typical scenario that visitors encounter at many science museum exhibits: an interactive feature and an interpretative label.

However, considering that none of the 52 children younger than age 13 were successful in the performance tasks, the conclusions of the author seem to ignore a major issue. Children and non-college educated adults are an important audience of museums and the exhibit was not effective in facilitating their science learning as measured by the performance tasks. The possible reasons for this abound, but one explanation may lie in the author's choice of exhibit topic. It may not have been reasonable to expect these concrete thinkers to master abstract light concepts in one short visit.

Designing exhibits for families

The series of studies of conducted by Borun and colleagues used observational and interview data to support design and evaluation efforts to promote family learning at four informal science learning institutions: two science museums, an aquarium, and a zoo (Borun et al., 1996; Borun, Chambers, et al., 1997; Borun & Dritsas, 1997). In the first study (Borun et al., 1996), three learning levels that progressed from less complex to more complex understanding were developed for each exhibit's goals at each institution. The learning levels, as expressed in interviews with 129 families, were compared with observed behaviors at exhibits to identify behaviors indicative of higher levels of learning. Five such "learning behaviors" were identified including: asking a question, answering a question, commenting or explaining the exhibit, reading text silently, or reading text aloud.

The second and third studies investigated which exhibit features promoted visitor use of these learning behaviors (Borun, Chambers, et al., 1997; Borun & Dritsas, 1997). Test exhibits were designed and redesigned at each institution using an iterative process to maximize family learning as indicated by the use of desired learning behaviors. However, the families were not interviewed to assess actual learning levels during these studies. Results of these studies found that the most effective exhibits for promoting desired learning behaviors among families were multi-sided (i.e., exhibits included three dimensional components), multi-user (i.e., multiple users could interact with exhibit components at one time), multi-outcome (i.e., exhibits allowed visitors to pick and choose which exhibit content to explore), multi-modal (i.e. exhibit components invited visitors to use different learning modalities: visual, auditory, tactile, or kinesthetic), readable, and relevant. In sum, these researchers found that such exhibits allowed family visitors to interact with the exhibits and each other in a manner that was more suitable for their individual needs and interests.

Designing exhibits for conceptual coherence

Falk (1997) used pre- and post-visit interviews to determine the influence of explicit labeling on the conceptual understanding of a purposeful sample of 174 museum visitors. The purposeful sample included 44% young adolescents (ages 11-15) and 56% adults (age 20+). The majority (81%) of participants attended one of two science museum exhibits with their families. About half of the visitors viewed the exhibits with explicit labeling while the rest viewed the exhibits without the explicit labeling. The explicit labeling consisted of a headline stating the main message of the exhibit and a sub-headline stating the main message of each exhibit element. The headline and sub-headline were posted at each exhibit element on temporary labels. Visitors from both age groups showed similar increases in conceptual understanding following their visit to the first exhibit on transportation regardless of exhibit labeling. In contrast, visitors from both age groups only showed increased conceptual understanding following their visit to the second exhibit on vertebrate development when the explicit labels were present. In this case, explicit labels seemed to cue visitors to important aspects of the exhibit they may not have perceived otherwise.

Findings of the preceding studies are instrumental for understanding how museum exhibits can and do facilitate the science learning of visitors. These studies rely on a variety of sources of evidence to support their claims of learning at exhibits. Museum exhibits that support the learning of many types of visitors encourage discussion and interaction. However, one study suggests that museum professionals need to be mindful of the needs and interests of visitors when selecting exhibits for design initiatives (Allen, 1997). Some exhibit content may not be suitable for all visitors due to visitor characteristics such as cognitive developmental level. If this is the case, design and evaluation efforts may be hindered due to validity problems.

Science Learning Experiences of Young Children

Young children learn about science through naturalistic, informal, and structured learning experiences (Lind, 2005). Naturalistic experiences are those in which children spontaneously begin exploring their world. Most young children's naturalistic experiences take place as a part of everyday activities in the home and community. In informal experiences, children spontaneously begin to explore, but at some point an adult intervenes to scaffold their learning. These experiences also often occur as a part of everyday activities. The adult in question is likely a parent or other caregiver. Both naturalistic and informal experiences are initiated by the child and are not planned. In contrast, an adult initiates and directs children's exploration in structured experiences. These experiences are planned and are likely to occur as part of preschool or kindergarten. In the case of structured experiences, the adult is usually a teacher or teacher's assistant.

Play

Play is the starting point for many young children's naturalistic and informal science learning experiences (Lind, 2005) and is essential for early social and cognitive development (Bredekamp & Rosegrant, 1995; Shonkoff & Phillips, 2000). The knowledge that children gain from these experiences has been termed "children's science" (Gilbert, Osborne, & Fensham, 1982). Many studies have documented children's science and its relationship to a wide range of science disciplines (Driver, Asoko, Leach, Scott, & Mortimer, 1994). However, the existing research documenting young children's naturalistic and informal science learning experiences that lead to these understandings is limited. A comprehensive search of major peer-reviewed journals in science education, environmental education, family studies, and leisure studies uncovered only two studies describing the science learning experiences of young children in the home. The vast majority of studies located describe the science learning experiences of young children in informal science learning environments, particularly museums.

Young Children at Home

Johnson, Alexander, Spencer, Leibham and Neitzel (2004) conducted a short longitudinal study of 211 four-year-olds residing in a rural university town to identify factors influencing the children's development of interests within conceptual domains. The families of the children volunteered for the study and were from higher socioeconomic backgrounds and were not ethnically or linguistically diverse. The parents completed initial questionnaires about the following areas: home environment and parental attitudes, socioeconomic status, and play behaviors. The children completed pre-tests of their cognitive abilities at the first meeting. Following this data collection session, parents were contacted by phone at two and four months to report on their children's play activities. Families of children with sustained interests in a conceptual domain (i.e., area of science) received a home visit and the children completed a post-assessment of their domain-specific knowledge.

From the parents' reports of their children's play, interest profiles were then developed to categorize the children's sustained interests. Over 40% of the boys exhibited sustained interests within conceptual domains while less than 13% of girls demonstrated such interests. The majority of these interests were related to the life and physical sciences. A comparison between the interest support activities of parents of children with sustained conceptual domain interests and parents of children with sustained non-conceptual domain interests (e.g., crafts, sports) revealed that the parents of children with conceptual domain interests reported greater involvement in activities related to obtaining information, such as reading non-fiction science books. Furthermore, nearly 70% of the parents of children with conceptual domain interests reported taking trips to informal learning settings that aligned with the children's interests while only 25% of parents of children with non-conceptual domain interests did so. Examples of such trips included museum and zoo visits.

More significantly, this study found that nearly 70% of children with sustained interests of any kind lived with another person who shared the same or similar interest, suggesting that deep interests in a topic are learned. However, the sample of families in this study represents a higher socioeconomic status than the general population and previous research has shown that higher socioeconomic groups typically participate in educational trips and activities more frequently than the population at large (Burkham, Ready, Lee, & LoGerfo, 2003; Hood, 2004).

Korpan, Bisanz, Bisanz, Boehme, and Lynch (1997) conducted interviews with 29 mothers of kindergarteners in a large city in Canada to determine the science learning experiences of young children. The researchers solicited volunteers from the schools to participate and only mothers responded. No documentation of socioeconomic status of the individual families was provided; however, the schools of the participating children were located in a middle class neighborhood. The majority of mothers (79%) reported that their children watched science-related television programs and did so an average of three times per week. The authors noted that many of the television programs described by the mothers were not specifically designed to be viewed by young children.

In addition, the majority of mothers (83%) participating in this study reported reading texts about science, nature, and technology topics to their children. Out of an average of over eight shared reading occasions per week, mothers reported that they read science, nature, and technology-related texts to their children more than a third of the time. This amounts to an average of about three shared readings of science, nature and technology-related texts per week, a similar frequency compared to the children's viewing of science-related television programs mentioned earlier. Such science, nature, and technology-related texts were varied and included children's fiction (83%), children's non-fiction (90%), children's magazines (72%), other fiction books (3%), other non-fiction books (45%), and other magazines and newspapers (48%). Interestingly, no mothers reported that their children used computers to learn about science, although 41% of the families had a computer at home and another 25% accessed a computer elsewhere. This may be because of the relative newness of home computer and Internet use at the time this study was conducted and the young age of the children.

Apart from media use, most of the mothers (86%) in Korpan et al.'s study (1997) also reported that their children participated in science-related observations and experiments at home such as mixing colors, watching plants and animals grow, and sky-watching. Additionally, nearly all mothers reported that their children participated in community activities pertaining to science, nature, and technology such as visiting zoos and museums and attending science courses at various facilities. The mothers reported that their children participated in such activities an average of 12 times per year. Although the socioeconomic status of the families in this study was not documented, it is likely the families belonged to a more privileged socioeconomic group and thus were able to provide their children with a variety of informal science learning opportunities. Furthermore, the location of the study in a large city likely influenced the number of science learning opportunities available to these families. Cities often have more museums, zoos, and other informal learning settings compared to rural areas. Lastly, due to the use of interview data in this study, there may be some inflated self-reporting bias, especially related to the type or frequency of activities described.

The paucity of available research documenting how young children learn science at home leaves many questions unanswered. However, the little evidence that does exist suggests that the home environment does play a role in the science learning of young children. Parents provide a variety of materials and experiences to support their children’s developing interest in and knowledge of science (Crowley & Jacobs, 2002; Palmquist & Crowley, 2007).

Young Children in Informal Science Learning Environments

By and large, the literature on early informal science learning experiences focuses on children learning at science museum exhibits with their families. More specifically, this line of research focuses on how parents scaffold the learning of their children at science museum exhibits. Researchers have concluded that parents employ a wide variety of scaffolding strategies to support their children’s learning at museum exhibits. They have also found that parents appear to use these strategies differently with boys and girls and with novice and expert children. These conclusions are mostly based on observations of family conversations and behaviors at exhibits.

Parent scaffolding at exhibits

Observations of families at exhibits have revealed that parents both help their children interact with exhibit features and guide their understanding of exhibit content (Allen, 2002; Ash, 2003, 2004, 2007; Blud, 1990a; Borun et al., 1996; Borun, Chambers, et al., 1997; Borun & Dritsas, 1997; Brown, 1995; Chrispeels, 1996; Crowley, Callanan, Jipson, et al., 2001; Fender & Crowley, 2007; Gleason & Schauble, 1999; Kelemen, Callanan, Casler, & Pérez-Granados, 2005; Palmquist & Crowley, 2007; Schauble & Bartlett, 1997; Siegel, Esterly, Callanan, Wright, & Navarro, 2007; Valle & Callanan, 2006; Zimmerman, Reeve, & Bell, 2009). Young children in particular are still developing some fine motor skills and need help manipulating scientific equipment and tools (Lind, 2005). They also are unfamiliar with the many roles adults adopt in everyday life. Shine and Acosta (2000) described instances in which 30 parents showed their young children (ages 4-6) how to use props and set the scene for role plays at facsimiles of a grocery store, doctor's office, and an ambulance in a children's museum. Schauble and Bartlett (1997) noted the importance of these behaviors for children understanding "what you are supposed to do" at exhibits in their description of designing a science gallery for families.

Parent directions

Similarly, Crowley and colleagues noted that many parents offered directions for using a zoetrope, a device that demonstrates the illusion of motion, or modeled how to use it in their observations of 91 families with young children (ages 4-8) at an animation exhibit in a science museum (Crowley, Callanan, Jipson, et al., 2001). By helping their children understand how to use the zoetrope, parents may have allowed their children to move beyond learning how to simply use the device (i.e.,, procedural knowledge) to exploring how it produced the illusion of motion (i.e.,, conceptual knowledge). However, it is still uncertain whether this assistance enhanced children's learning because no direct measurements of child learning were included in this study.

Parent explanations

In addition to showing children how to use equipment in museum exhibits, parents often assume the role of explainers at exhibits, offering simple explanations about the phenomena demonstrated (Borun et al., 1996; Borun, Chambers, et al., 1997; Borun & Dritsas, 1997; Crowley, Callanan, Jipson, et al., 2001; Crowley, Callanan, Tenenbaum, et al., 2001; Fender & Crowley, 2007; Gleason & Schauble, 1999; Kelemen et al., 2005; Palmquist & Crowley, 2007; Siegel et al., 2007; Valle & Callanan, 2006). Crowley, Callanan, Tenenbaum, et al. (2001) called these explanations "explanatoids" and describe them thusly: "simple, incomplete, and mundane-no more than a few words uttered by a parent at the appropriate moment during the ongoing activity" (p.260). These explanations of phenomena may be based on exhibit labels or relevant prior experiences. They may also include the use of comparisons between phenomena and analogies about the relationships between phenomena.

In a study of 64 families at the same zoetrope exhibit, Fender and Crowley (2007) documented that young children (ages 3-8) who heard parent explanations were more apt to correctly describe the function of a zoetrope and relate it to other animation devices in a post-test than children who used the device by themselves or children whose parents did not offer explanations. Thus, children who heard parent explanations demonstrated conceptual, rather than just procedural, understanding of the zoetrope device.

Parent assistance in problem-solving and scientific reasoning

During museum visits, parents scaffold children's problem-solving and scientific reasoning by focusing their attention on important evidence, directing them to make productive comparisons, and generally guiding their thinking about evidence (Crowley, Callanan, Jipson, et al., 2001; Crowley, Callanan, Tenenbaum, et al., 2001; Gleason & Schauble, 1999; Siegel et al., 2007). The amount of direction parents provide their children in conducting problem-solving and inquiry activities during museum visits seems to differ by parental education levels and the children's ages.

In a study of 40 Mexican-descent families, Siegal, Esterly, Callanan, Wright, and Navarro (2007) observed that parents with higher education levels tended to take a more directive approach, in which they behaved as experts conveying directions and information to the children, during both a sink-and-float investigation and explorations of other exhibits at a children's museum. These parents may be more accustomed to being schooled in the traditional way (i.e., direct instruction), which they apparently adopt to teach their own children. Parents of younger children (aged three to five years) also demonstrated a more directive approach, which Ash (2004) suggests may be indicative of a greater focus on language learning rather than science learning when visiting museums. Parents of young children seem to intuitively understand the axiom, "language before learning," and spend a great deal of time in museums modeling the use of language as they name and describe objects. Again, it is unknown whether a parent's interaction style influences children's learning because no direct measurement of children's learning was included in this study.

Differences in parent scaffolding at exhibits

Some evidence suggests that parents may not interact with boys and girls equally at museum exhibits (Blud, 1990b; Crowley, Callanan, Tenenbaum, et al., 2001). Crowley, Callanan, Tenenbaum, et al. (2001) examined the conversations of 298 families with at least one young child (aged one to eight years) at exhibits at a children's museum and found that parents were three times more likely to offer explanations to boys than to girls. However, the incidence of parent-child discussions about how to use exhibits and evidence from the exhibits was similar for boys and girls. Interestingly, the boys did not solicit these explanations any more than the girls by asking questions, yet the gender difference was observed for even the youngest group of children (aged one to three years). Given that the learning of children was not directly assessed in this study, it is still uncertain whether this gender difference in parent explanations contributes to a gender difference in children's science learning at exhibits.

Additional evidence suggests that parents may not interact with expert and novice children equally. Palmquist and Crowley (2007) examined the conversations of 42 families with at least one young child aged five to seven years at a dinosaur exhibit at a natural history museum. In addition, the researchers administered a pre-assessment of the children's prior knowledge about dinosaurs and paleontology and a parent survey. The parent survey collected information in several areas including the parents' self-reported interest in and knowledge of dinosaurs, their child's perceived interest in and knowledge of dinosaurs, family museum attendance, and the child's favorite toys, books, and activities, especially those related to dinosaurs.

The results of the parent survey indicated that the children classified as dinosaur "experts" were largely male and had families that shared and supported their interest in dinosaurs. A comparison of conversations of families with dinosaur experts and dinosaur novices revealed that parents of novices more actively engaged their children in conversation. More specifically, parents of novices took more turns talking to their children than parents of experts. Parents of expert children were relatively silent and did little to encourage their children to expand their understanding of dinosaurs. No post-test was used in this study and so it is unknown how much the novice or the expert children were able to advance their understanding of dinosaurs as a result of visiting the exhibit.

Although research into the behavior of families at exhibits suggest that parents support their children’s learning in many ways, few studies have documented how parent scaffolding impacts the learning of young children (Fender & Crowley, 2007). Instead, researchers have inferred that children learn more in the presence of their parents because they spend more time in focused exploration or hear helpful hints about using or understanding exhibits. Additionally, research has documented differences in parent scaffolding based on a child’s gender and experience and the parent’s educational level. Even though these differences are of great concern to science educators, the implications of these findings are not supported by data indicating the impacts of these differences on children’s learning.

What Young Children Know about Living Things

From their earliest learning experiences, young children classify living things based on the characteristic of movement (Hatano & Inagaki, 1995; Venville, 2004). Venville (2004) observed that 26 young English children (aged five to six years) attributed a variety of other characteristics to living things such as the ability to die, grow, reproduce, sense, drink, eat, and make noise. In this study, some children also characterized living things as possessing certain body parts (e.g., legs, face, and hearts) or resembling humans in some manner. Depending on the criteria the children used to classify things as living or nonliving, some children classified nonliving things as living such as the sun and fire and living things as nonliving such as plants.

Classification of Living Things

Physical characteristics are the primary way young children classify living things (Barrow, 2002; Prokop, Prokop, & Tunnicliffe, 2007; Shepardson, 2002; Trowbridge & Mintzes, 1988; Tunnicliffe & Reiss, 1999). Young children classify animals in particular using external anatomical features such as legs and wings (Barrow, 2002; Prokop et al., 2007; Shepardson, 2002). Prokop et al. (2007) found that 445 young Slovakian children (aged six to eight years) showed little understanding of the internal anatomy of invertebrates and often drew bones and lungs inside invertebrates in their drawings of stag beetles and crayfish.

Insects

Barrow (2002) interviewed 24 primary-aged American children about the characteristics of insects and found that they mentioned a variety of characteristics including small size, eyes, six legs, four wings, and antennae. Small size was the most common feature mentioned. However, none of the children noted the jointed legs or exoskeleton of insects and only one child talked about the three-part body of insects. Shepardson (2002) interviewed 60 primary-aged American children about their drawings of insects and how they classified pictures of various insects. The children used the characteristics of small size, bug-like shape, legs, and antennae in their explanations of what makes living things insects. Furthermore, the insect understanding of children seemed to be linked to “prototypical insects” such as beetles and butterflies, which were the most common insects in a study of 218 English children’s drawings of their favorite insect (Snaddon & Turner, 2007). However, spiders (which are arthropods but not insects) were drawn by many of the young children in both the Shepardson (2002) and Snaddon and Turner (2007) studies.

Shepardson (1997) and Barrow (2002) found that children had limited understanding of insect life cycles prior to instruction. The 24 first-grade children in Shepardson’s study often represented the larval, pupal, and adult stages of the butterfly life cycle but not the egg stage in their journal entries and interviews. In contrast, the children in Barrow’s study omitted the pupal stage instead of the egg stage. Shepardson further noted that some of the children believed that all larvae became butterflies or moths, when in fact there are many other groups of insects that pupate. Other children equated larvae such as mealworms with earthworms due to their similar characteristics such as body shape and movement.

Interdependency of Living Things

Several researchers have concluded that young children have limited understanding of the interdependency of living things in general (Gallegos, 1994; Leach, Driver, Scott, & Wood-Robinson, 1996b; Myers, Saunders, & Garrett, 2004; O’Byrne, 2008; Snaddon, Turner, & Foster, 2008; Strommen, 1995), and the ecological role of insects in particular (Shepardson, 2002). Myers et al. (2004) used interviews and drawings to elicit 141 young zoo visitors’ understanding of the needs of a favorite animal. The children, aged four to fourteen years, almost exclusively drew vertebrates with mammals accounting for nearly 70% of the animals drawn. Invertebrates accounted for less than one percent of the children’s favorite animals.

Although not discussed in the Myers et al. (2004) study, the vertebrate focus of children may be influenced by their limited understanding of the classification of animals (Trowbridge & Mintzes, 1988). In addition, Myers et al. found that many young children (aged four to seven years) were aware of many of the physiological needs of animals such as air, food, and water but few were aware of any ecological needs such as habitat or relations with other species. More specifically, Shepardson (2002) observed that young children “think about what organisms eat in a one-dimensional animal-food (insect-food) relationship that considers what an insect eats but not what eats the insect” (p.639). The most common understanding of the relationship between insects and humans among the children studied was the idea that insects were harmful to people. Young children in particular discussed stinging and biting insects.

Research documenting the understanding young children have about living things, animals (particularly insects), and ecology suggests that young children have at least a partial framework for learning about these topics that should be considered when designing formal and informal learning experiences. Young children are able to identify some of the characteristics of living things and classify animals such as insects based on observable physical features such as legs and wings. However, their understanding of the characteristics and classification of living things may not match the accepted scientific conceptions of these topics. It seems that young children have limited first-hand experiences with many living things, including invertebrates such as insects. Insect-related topics that appear difficult for young children to conceptualize are the life cycle of insects and the role of insects in ecosystems. Questions remain regarding whether these difficulties in understanding are related to the abstract nature of the content or to the limited experiences young children have with invertebrates, insects in particular.

Assessment of Visitor Learning in Museums

Early research on learning in museums relied on paper-and-pencil tests using quantitative approaches strongly influenced by traditional classroom assessment practices. In these early studies, visitors, some even as young as seven years old, completed pre- and post-tests of their content knowledge (Falk, 1982; Melton, 1936). These tests often included multiple-choice questions that constrained visitor responses but were easily scored and analyzed. The tests targeted specific content that all visitors, regardless of age or ability, were expected to learn during a visit. In this way, museum visits were viewed as analogous to classroom lessons.

Few recent studies of cognitive learning in museums have employed paper-and-pencil tests of content knowledge (Anderson & Lucas, 1997; Kubota & Olstad, 1991). Instead, most recent studies have used more open-ended qualitative assessments (e.g., interviews, drawings, concept mapping) in order to capture the full range of visitor understanding. Observational studies have also been instrumental in assessing visitor learning at museum exhibits. These open-ended approaches are more aligned with recent position statements regarding the most effective research and evaluation approaches for studying informal science learning environments (Allen, 2008; Dierking et al., 2003).

Interviews

Many museum researchers have used interviews for pre- or post-assessments of visitor knowledge (Allen, 1997, 2004; Borun et al., 1996; Cox-Petersen et al., 2003; Ellenbogen, 2002; Falk, 1997; Falk & Adelman, 2003; Packer & Ballantyne, 2005;

Palmquist & Crowley, 2007). Interview approaches vary but generally, they are semi-structured and take place face-to-face at the museum. Both closed and open-ended questions about exhibit content have been asked (Falk, 1997). Furthermore, interviews have incorporated props, photographs, or maps to stimulate visitor thinking about specific exhibit content (Allen, 1997; Allen, 2004; Palmquist & Crowley, 2007). Interviews allow visitors of many different ages and abilities to communicate their understanding. However, interviews are more difficult to score reliably and researchers using interviews have provided limited evidence of the reliability or validity of the interview protocols used.

Observations

Many researchers have also used observations to assess visitor learning at museum exhibits (Allen, 2002; Ash, 2003; 2004; 2007; Borun et al., 1996; Borun, Chambers, et al., 1997; Borun & Dritsas, 1997; Chermayeff et al., 2001; Crowley, Callanan, Jipson, et al., 2001; Crowley, Callanan, Tenenbaum, et al., 2001; Ellenbogen, 2002; Gilbert & Priest, 1997; Gleason & Schauble, 1999; Palmquist & Crowley, 2007; Tunnicliffe, 2000; Tunnicliffe et al., 1997). Researchers have recorded and analyzed conversations and behaviors of visitors at exhibits in order to infer that learning has occurred. A variety of analysis techniques have been used to quantify this qualitative data. For example, Tunnicliffe (2000) and Tunnicliffe et al. (1997) used a “systematic network” to organize and analyze the content of complete conversations between adults and children at exhibits. In contrast, Ash (2003) looked at both the content and inquiry skills present in representative dialogic segments, which were specific instances when families discussed content and used inquiry skills at exhibits. The studies incorporating behavioral data examined how parents support their children’s interactions with exhibits

(Borun et al., 1996; Borun, Chambers, et al., 1997; Borun & Dritsas, 1997; Crowley, Callanan, Jipson, et al., 2001; Crowley, Callanan, Tenenbaum, et al., 2001; Gleason & Schauble, 1999).

In families in particular, a wide variety of behaviors have been documented and analyzed for their potential in advancing children's exploration of exhibits and understanding of exhibit content. Observations of visitors give researchers a picture of visitor learning that is contextualized and free of biases associated with traditional assessment measures. However, inferences about learning based on observations are stronger when additional forms of evidence such as interview data are used. Only Borun and colleagues have suggested a method of evaluating visitor conversations and behaviors at exhibits, but limited reliability and validity evidence has been presented for these studies (Borun et al., 1996; Borun, Chambers, et al., 1997; Borun & Dritsas, 1997).

Drawings

Researchers working with younger museum visitors have used drawings to aid recall during interviews (Myers et al., 2004; Piscitelli & Anderson, 2001). Piscitelli and Anderson (2001) asked children to draw any aspect of museums they wished to while Myers et al. (2004) asked children to draw their favorite animal in the zoo. The researchers referred to the drawings during subsequent interviews with the children in order to stimulate conversations about their experiences and understandings. The researchers also completed basic content analyses of the features of the children's drawings. Drawings can provide researchers with insights regarding what young visitors think about their experiences and understandings, but as is the case with observations, drawings are more valuable for assessment purposes when they are coupled with additional data sources regarding learning. Interestingly, no prominent museum researchers have suggested or developed valid and reliable methods of scoring drawings.

Personal Meaning Mapping

Researchers striving to gain insights into how visitors think and feel about museum exhibit content have used a modified version of concept mapping called personal meaning mapping (Anderson et al., 2000; Falk et al., 1998; Falk & Storksdieck, 2005; Storksdieck, Ellenbogen, & Heimlich, 2005). Personal meaning mapping is “designed to measure how a specified ‘educational’ experience uniquely affects each individual's personal conceptual, attitudinal, and emotional understanding (Falk et al., 1998, p. 108). Despite this attitudinal focus, the scoring method described by Falk (1998) seems to target cognitive outcomes rather than affective outcomes. In his study, visitor maps produced before and after viewing museum exhibits were scored using four criteria: use of appropriate vocabulary, depth of conceptual understanding, breadth of conceptual understanding, and mastery of a topic. In addition, the discussions in Falk (1998) and Falk and Storksdieck (2003) did not mention affective outcomes related to visitor experiences at exhibits.

Perhaps recognizing that personal meaning mapping may not represent changes in attitudes as well as changes in knowledge, Storksdieck et al. (2005) incorporated a card sort task in addition to personal meaning mapping in their study. However, the visitors in their study interacted with an exhibit about biodiversity and conservation. The content of this exhibit likely elicited more affective comments during the personal meaning mapping sessions than the content of the exhibits in the other two studies, gems and minerals and characteristics of life respectively. These studies have shown that raters can score personal meaning maps reliably; however, the claim that personal meaning maps provide information about affective outcomes is not well substantiated. It may be that the ability of personal meaning mapping to provide evidence related to affective outcomes is related to the content of the exhibit.

Observations and interviews are currently the most common methods used to assess learning in museum settings due to their ability to capture a broader range of visitor understandings. In addition, unobtrusive observations generally do not influence the learning of visitors at exhibits, but insights gained from such observations are more useful when supported with additional evidence. On the other hand, pre-visit interviews may cue visitors to pay more attention to certain features or exhibit content but can provide strong evidence of learning. More consideration of reliability and validity issues associated with different methods for scoring interview and observation data is needed. Additional methods of documenting the learning of museum visitors should be developed that respect the varied understandings of visitors and can be used to complement observational data.

Assessment of Young Learners in Science Education

Assessment of young children often begins with observations or “child-watching” (Mindes, 2003). Practitioners may make informal observations of children playing, manipulating objects, and interacting with their peers. Alternatively, practitioners may record observations, evaluate samples of children’s work, or interview children in more formal assessments. Britsch (2001) recommended that both oral and written assessment formats be used to assess the “emergent science literacy” of young children.

Combining Assessment Types

Accordingly, several researchers have used paper-and-pencil tests and interviews to gather information about seven and eight-year-olds' understandings of ecological concepts (Leach, Driver, Scott, & Wood-Robinson, 1995, 1996a, 1996b; O'Byrne, 2008). Paper-and-pencil tests have limited efficacy in assessing the conceptual understanding of young children because children at this age may not be able to read test items or explain their ideas (Siegal, 1999). However, such tests may be useful as a tool to activate young children's prior knowledge prior to follow-up interviews (O'Byrne, 2008). Other researchers have combined drawing and interview data to measure young children's understanding of animal classification and life cycle concepts (Prokop et al., 2007; Shepardson, 1997; Shepardson, 2002). This approach is similar to the interview process used by Gilbert et al. (1982) to document children's conceptual understanding related to a variety of science topics.

Concept Mapping

Concept mapping is another assessment tool commonly used to probe children's conceptual understanding of science (Novak & Gowin, 1984). Concept mapping has customarily been used with older children who can read and write although some researchers have begun to assess the conceptual understanding of younger children with modified versions (Figueiredo et al., 2004; Hunter et al., 2008; Monroe-Ossi, Wehry, Algina, & Hunter, 2008). Gallenstein (2003) recommended that the format for concept maps should match children's developmental levels and thus young children should use objects or pictures more than written words. Accordingly, Figueiredo et al. (2004), Hunter et al. (2008), and Monroe-Ossi et al. (2008) asked the preschool children in their studies to create concept maps depicting their understanding of life science concepts using pictures. The children described the relationship between the pictures orally and the researchers recorded their thoughts on sticky notes.

The assessment of young children in science education must take into account their literacy and developmental levels. More specifically, young children may not be able to read or write proficiently enough to permit the reasonable use of paper-and-pencil tests. At younger ages, assessments such as interviews and drawings better match their literacy skills and development levels. In addition, concept mapping tasks using pictures and concrete objects is beginning to be explored in science education as an alternative assessment strategy.

Concept Maps

Concept maps are graphical representations of the structure of a person’s declarative knowledge (Novak & Gowin, 1984). They are formed from a collection of propositions organized from general to more specific in a hierarchy or in a network (Novak & Gowin, 1984; Ruiz-Primo & Shavelson, 1996). Propositions are created when two related concepts are connected by a labeled line or arrow. The words on the labeled line or arrow are known as linking words and describe the relationship between the two concepts. The use of linking words allows each proposition to be read as a sentence. For example, the proposition in Figure 2-1 can be read as: Peanut butter and jelly is a type of sandwich.

Reliability and Validity of Concept Maps as Assessment Tools

Ruiz-Primo and Shavelson (1996) first issued the call for increased consideration of the reliability and validity of concept mapping tasks after noting that few researchers using concept maps as assessment tools addressed issues of reliability and validity comprehensively. Whereas reliability refers to the consistency of an assessment’s scores across multiple administrations, items, forms, and judges, validity refers to the appropriateness of an interpretation or use of such scores. In Ruiz-Primo and Shavelson's review of over 20 studies of concept maps as assessments, the most commonly-reported type of reliability evidence was inter-rater reliability while the most commonly-reported types of validity evidence were content and convergent evidence.

On reliability

Ruiz-Primo and Shavelson (1996) conceived of concept maps as a combination of "a task that invites students to provide evidence bearing on their knowledge structure in a domain, a format for the students' responses, and a scoring system by which students' concept maps can be evaluated accurately and consistently" (p. 573). They noted that an assessor's task, format, and evaluation choices all have implications for the reliability, or consistency, of concept map scores. This understanding prompted Ruiz-Primo, Shavelson, and their colleagues to undertake studies examining the reliability of different combinations of concept mapping tasks and scoring systems. The equivalence of different concept mapping formats, however, has yet to be examined.

Concept map tasks vary in regards to task demands, task constraints, and task content structures (Ruiz-Primo & Shavelson, 1996). Task demands refer to what individuals are asked to do, which can range from constructing a concept map to filling in a blank map. Alternatively, individuals may be asked to answer interview questions, sort cards, or rate relatedness of concept pair. Task constraints include the requirements and limitations placed on individuals regarding how they complete the task. For example, individuals may be required to construct a map with a hierarchical structure or concepts and linking phrases provided by the assessor. Task content structures are the possible interactions between the task demands, the task constraints and the structure of the subject domain to be mapped. While Novak & Gowin (1984) proposed that all concept maps should be hierarchical, Ruiz-Primo and Shavelson (1996) argued that the map structure should match the content structure. In other words, task constraint choices should be sensitive to the content structure.

Researchers comparing the reliability of concept map scores from different tasks have distinguished between tasks that are more directed and tasks that are less directed. More directed tasks include fill-in-the-map tasks where the assessor provides the structure, the concepts, and sometimes even the linking terms. In contrast, less directed tasks include construct-the-map tasks where the assessor may provide the concepts but nothing else. Construct-the-map tasks where the assessor provides some or all of the concepts would be considered more directed than construct-the-map tasks where the individual chooses the concepts. In general, researchers have found that assessor choices regarding concept mapping tasks introduce error into map scores and that more directed tasks have higher inter-rater reliability estimates (Ruiz-Primo, Shavelson, Li, & Schultz, 2001; Shavelson & Ruiz-Primo, 2000; Yin & Shavelson, 2008; Yin, Vanides, Ruiz-Primo, Ayala, & Shavelson, 2005).

For example, Ruiz-Primo et al. (2001) conducted a generalizability study with 43 high school chemistry students in two intact classes. The students constructed concept maps about ions, molecules, and compounds. A fully-crossed study design was employed in which students were crossed with tasks and raters. Each student was evaluated on three tasks: construct-the-map with one set of assessor-provided concepts, construct-the map with a second set of assessor-provided concepts, and construct-the map with no concepts provided. All students constructed their first map using their own concepts and they constructed their second and third maps with two lists of concepts which were generated by randomly selected concepts from a master list developed by content experts. The order in which students received the lists was randomly counterbalanced. Furthermore, student maps were evaluated by two different raters using three different scoring systems.

Map scores were analyzed using generalizability theory. The largest variance components from the generalizability studies were for nonrandom differences between persons followed by the component for the person-by-task condition (Ruiz-Primo et al. 2001). This finding suggests that concept maps are able to detect systematic differences in students' knowledge of content. However, the task condition did introduce some variability in the students' relative rankings for the three scores. More specifically some students performed better with a set of concepts provided while others performed better with their own concepts. As Ruiz-Primo et al. pointed out, task sampling often introduces error into the scores of performance assessments such as concept maps.

In contrast, the variance components for raters and all interactions with raters were negligible. Raters were able to score the students' maps consistently. Further analysis revealed that the inter-rater reliability estimates for the maps using the assessor-provided concepts were higher overall than maps using student-generated concepts. Again, Ruiz-Primo et al. noted that this is typical of performance assessments, as raters have been shown to be capable of making complex scoring decisions consistently.

Individuals may respond to concept map tasks in a variety of ways that differ by response mode, characteristics of the response format, and the mapper (Ruiz-Primo &

Shavelson, 1996). Response modes include paper-and-pencil, oral, and computer-generated. The most common response mode described in the literature is paper-and-pencil where individuals draw a map on paper or write in concepts or linking words in blank spaces on a skeleton map. The characteristics of the response format can affect how individuals represent their understanding. For instance, a skeleton map with a hierarchical structure essentially forces all individuals to represent their understanding hierarchically. Last, the mapper may be the individual or the assessor. Assessors may develop concept maps from written essays or oral responses (Novak, 2005). Researchers have not yet compared the reliability of different response formats.

Researchers have developed a variety of scoring systems that evaluate the selected components of concept maps, including propositions, hierarchy levels, and examples. Novak and Gowin (1984) included all map components in their scoring system and awarded points for each valid proposition, level of hierarchy, example, and cross-link. Because they assumed that the structure of knowledge is hierarchical (Ausubel, 1960), they placed greater emphasis on the use of multiple levels of hierarchy and cross-links and so awarded a greater number of points for these components. Consequentially, Novak and Gowin’s scoring system is only appropriate for use with hierarchical concept maps.

Alternatively, a number of researchers have employed propositional scoring systems in their work. Propositional scoring systems can be used with concept maps containing a variety of structures because they focus only on the propositions in a map. Although many researchers have failed to articulate any cognitive theory as a basis for their choice of a propositional scoring system, network theory is implied because of the emphasis on the quality of the connections between concepts in a propositional scoring system. In one example, McClure and Bell (1990) awarded a point for each part of a proposition (e.g., concepts, arrows, and linking words) that validly represented the relationship between two concepts in a proposition. In another example, Ruiz-Primo et al. (2001) and Shavelson and Ruiz-Primo (2000) judged each proposition on a five-point scale from zero points for inaccurate/incorrect to four points for excellent/outstanding. In a last example, Klein, Chung, Osmundson, Herl, and O'Neil (2002) awarded half a point for propositions that appeared in one of two expert maps and a full point for propositions that appeared in both expert maps.

Researchers have explored the tension between efficiency and reliability in the use of different concept map scoring systems. Efficiency is related to the amount of time and cognitive effort different scoring systems demand. Some scoring systems require complex decisions to be made about many different map components and thus require more time and cognitive effort to use. Researchers have investigated different scoring systems in an attempt to identify the most efficient and reliable system for scoring concept maps. Such knowledge would be beneficial to both researchers and classroom practitioners using concept maps as assessments.

In general, researchers have found that more efficient scoring systems have similar, if not better, reliability indices compared to less efficient scoring systems (Klein et al., 2002; McClure, Sonak, & Suen, 1999; Shavelson & Ruiz-Primo, 2000). For instance, McClure et al. (1999) used six different scoring methods to score a set of concept maps produced by 63 undergraduate students enrolled in an introductory educational psychology course. The students constructed the maps using a list of 20 concepts provided to them. The entire set of maps was scored by six pairs of independent raters with each pair using a different method: holistic, holistic with master map, propositional, propositional with master map, structural, and structural with master map. The holistic method was developed specifically for use in this study while the propositional and structural methods were based on the previous work of McClure and Bell (1990) and Novak and Gowin (1984), respectively. The master map was developed by the professor teaching the course and was provided to the raters as “a guide in the application of the assigned scoring method” (McClure et al., 1999, p.484). The authors employed a fully crossed design in which students were crossed with raters.

The results of a generalizability study indicated that the scoring method used greatly impacted the reliability of scores. The most reliable scores were obtained using the propositional scoring method with master map. In addition, the authors noted that this method required raters to make few complex decisions and spend a moderate amount of time scoring maps compared to the other methods. Thus, the authors recommended the use of the propositional scoring method with master map to others.

Other concept map studies have generally reported the reliability of concept map scores derived from a single task, format, and scoring system instead of making comparisons (Rice, Ryan, & Samson, 1998; Rye & Rubba, 1998; Schau, Mattern, Zeilik, Teague, & Weber, 2001; Stoddart, Abrams, Gasper, & Canaday, 2000; Van Zele, Lenaerts, & Wieme, 2004). All but one of the studies listed above only reported inter-rater reliability of concept map scores while the remaining study reported only the internal consistency of concept map scores (Schau et al., 2001). The reported inter-rater reliabilities in all of these studies were high.

These study designs used only one administration of the concept map task, which eliminated the possibility of determining the variability of scores due to occasion. Also, the types of tasks used were less directed, which made it difficult to determine variability in concept map scores due to task. In contrast, the task in the study where internal consistency was reported was very directed as it was a select-and-fill-in mapping task (Schau et al., 2001). More specifically, the middle school students in this study filled in concepts in skeleton maps using an assessor-generated list. Each concept could only be used once in the maps. Thus, each concept or item on the maps could easily be evaluated without concern for scoring variability and the item scores could be correlated using Cronbach's alpha. In sum, this suggests that the feasibility of looking at the internal consistency of concept map scores varies with the directedness of the task.

On validity

As is the case with reliability concerns, an assessor's choices of task, format, and scoring system have implications for the validity of score interpretations (Ruiz-Primo & Shavelson, 1996). The range of directedness of the tasks used has raised validity concerns among Ruiz-Primo and Shavelson (1996). These researchers have questioned whether different mapping techniques (e.g., construct-a-map, fill-in-a map) produce scores that represent similar aspects of individuals' knowledge structures. They contend that more directed tasks are easier to score and evaluate for reliability and thus are more attractive to stakeholders interested in using concept maps as alternative assessments in research and practice. However, some evidence suggests that the directedness of tasks may influence the way individuals respond in a variety of ways (Yin & Shavelson, 2008).

More directed tasks may unduly influence how individuals represent their understanding in concept maps. Along the same line of thinking, the development of multiple scoring systems has prompted some researchers to ask whether different scoring systems produce scores that represent individuals' knowledge structures equivalently. Researchers have conducted cognitive analyses and comparisons across measurement methods to validate their interpretations of concept map scores derived from different concept map tasks and scoring systems. As is the case with reliability, the validity of different concept map response formats has not been investigated.

Validity evidence based on content has been cited in several studies and is commonly described in the development of concept map tasks. Content-based evidence involves examining the relationship between an assessment's content and the content area it was designed to measure. Researchers have used many techniques to document this type of evidence in their studies. First, many have consulted with content experts in the development of assessor-provided lists of concepts and linking terms for some construct-a-map and fill-in-a-map tasks (Klein et al., 2002; Shavelson & Ruiz-Primo, 2000). Other researchers have asked content experts to generate their own concept maps which are then used to develop skeleton maps for fill-in-a-map tasks (Ruiz-Primo et al., 2001; Yin & Shavelson, 2008). Still others have used these expert maps as criterion maps which are then used for comparison in scoring procedures (Klein et al., 2002; McClure et al., 1999; Rye & Rubba, 1998). Lastly, some researchers have involved content experts in making decisions about the quality of propositions during the scoring process, especially when propositions are identified that confound the knowledge and expertise of raters (Stoddart et al., 2000). The input and judgment of experts in these studies ensures that the content of the maps is representative of the subject domain.

Additional content-based evidence for validity of concept mapping tasks can be found in the cognitive analyses some researchers have employed (Klein et al., 2002, Yin et al., 2005). Cognitive analyses such as written elaborations or “think-alouds” about propositions in a map can reveal not only the accuracy of content but also the depth and complexity of an individual’s understanding. In this way, cognitive analyses can uncover areas of deep understanding or profound misconceptions. Furthermore, cognitive analyses can pinpoint how the task limits an individual’s ability to express his/her understanding fully. In such a case, modification of the task may be warranted. Similarly, cognitive analyses can demonstrate how the scoring scheme underestimates an individual’s understanding. In this case, modification of the scoring scheme may be needed in order to better account for an individual’s understanding.

Many authors have also collected convergent evidence for the validity of their concept map score interpretations. Convergent evidence involves examining the relationship between the scores of measures targeting the same content. Studies have compared the performance of individuals on concept map assessments and at least one other assessment such as a multiple choice test, essay test, or card sort (Klein et al., 2002; McClure & Bell, 1990; Ruiz-Primo et al., 2001; Rice et al., 1998; Schau et al., 2001). Authors have used Pearson correlations and multi-trait, multi-method matrices to make these comparisons.

For instance, Klein et al. (2002) used a fully-crossed design in which each student’s understanding was evaluated for two concepts: the sense of hearing and the sense of vision using three different assessment methods: a multiple choice assessment, an essay assessment, and a concept map assessment. The concept mapping task asked students to construct a map using provided concepts and linking words. The response format was computer-generated and the scoring system awarded points for propositions matching expert-generated maps. Fifty-six fourth and fifth-grade students in two intact classes completed the three assessments at the end of the two instructional units on hearing and vision. The order in which the students completed the instructional units and the assessments was counterbalanced.

Klein et al. (2002) used a multi-trait, multi-method matrix to compare student performance on the assessments and found moderately positive correlations between the three measures. However, the correlations between the multiple choice and essay assessments were higher than the correlations between the concepts map and multiple choice assessments and the concept map and essay assessments. This finding suggests that concept maps measure overlapping, but not exactly the same, aspects of a domain of knowledge compared to multiple choice and essay assessments.

The finding that concept maps measure roughly the same aspects of a domain of knowledge is supported by the work of others using convergent validity evidence (McClure & Bell, 1990; Rice et al., 1998; Ruiz-Primo et al., 2001; Schau et al., 2001). These authors accept a moderately positive correlation as affirmation of their belief that concept maps better represent the range and diversity of an individual’s understanding compared to traditional assessment techniques. Construct-a-map tasks are viewed as particularly well-suited for this purpose even though the maps generated from these tasks are more difficult to score and use reliably. Construct-a-map tasks are preferred by some who cast doubt on the validity of scores generated from more directed concept mapping tasks. More directed concept mapping scores are thought to overestimate or constrain individuals' representations of their understanding (Ruiz-Primo et al., 2001, Yin et al., 2005; Yin & Shavelson, 2008). This creates a tension between reliability and validity issues when designing concept map tasks, response formats, and scoring systems.

Ruiz-Primo, Schultz, and Shavelson (1997) recommended that four criteria be considered when considering the use of concept maps as assessment tools: a) appropriateness of the cognitive demands required by the task; b) appropriateness of a structural representational in a content domain; c) appropriateness of the scoring system used to evaluate the accuracy of the representation; and d) practicality of the technique (Ruiz-Primo et al., 1997, p.7). These four criteria not only attempt to balance reliability and validity concerns, they also take into consideration feasibility concerns. Thus, the criteria provide guidance to researchers and practitioners interested in using concept maps to assess the learning of individuals of different ages and abilities in a variety of settings.

Summary

Among the numerous studies of families and children in museums, only a handful has focused on how museum experiences influence the learning of children. Young children in particular seem to be an invisible population in museum research and evaluation. Young children are far from invisible at museum exhibits as they are often the first in a group to touch and manipulate anything that can be touched or manipulated, sometimes quite vigorously. Anecdotal evidence suggests that museum experiences are enjoyable and memorable for young children (Piscitelli & Anderson, 2001). Observational studies of families with young children describe in great detail the ways parents scaffold their children’s interactions with exhibits (Crowley, Callanan, Jipson, et al., 2001). However, these studies leave many questions unanswered regarding the learning of young children at museum exhibits and the efficacy of parental scaffolding. Addressing this research gap is necessary if we want to understand the role early science learning experiences play in lifelong science learning and address any gender and socioeconomic differences associated with early science learning experiences (Crowley, Callanan, Tenenbaum, et al., 2001; Siegel et al., 2007).

Two main challenges must be overcome in order to begin addressing some of these questions and close the research gap. First, learning assessments of young children must take into consideration their literacy and developmental levels. Young children may not be able to understand or express themselves on traditional assessments. Interviews, drawings, and concept maps appear to be the most suitable methods for assessing the learning of young children in museum settings. Second, assessments of museum visitors must allow for a wide range of experiences and understandings. Interviews with at least some open-ended questions and concept mapping activities seem to be the most suitable methods for assessing the learning of museum visitors.

Concept mapping as an assessment tool allows for a wide variety of understanding to be communicated. In fact, it is widely accepted that no two concept maps are identical and many different concept maps can represent complete and complex understanding of a concept (Novak & Gowin, 1984). Previous research has demonstrated methods for using concept mapping reliably and validly. In addition, concept mapping can be modified for use with young children. Concept mapping formats for young children can eliminate the constraints of working with non-readers if pictures or concrete objects are used to represent concepts. Young children can manipulate these materials and explain their understanding of the relationships between concepts orally.

Based on the documented need for additional research in the specific areas discussed in this review of the literature, this research study aimed to investigate specific questions regarding the learning of young children on family visits to a live butterfly exhibit at a natural history museum by developing and using a concept mapping and “think-aloud” assessment format More specifically, this study identified the types of understandings about butterflies young children develop as a result of their experiences in a live butterfly exhibit. The study also identified differences in the understandings of children who have and have not recently visited the exhibit.

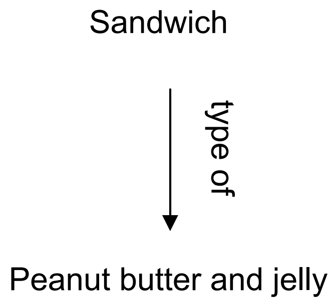


Figure 2-1. Example proposition

CHAPTER 3
METHODOLOGY

Introduction

The purposes of this study were to: 1) validate the use of a concept map assessment tool specifically designed for use with young children (aged five to seven years); and 2) use this tool to document similarities and differences in the learning of young experienced and inexperienced visitors in a unique setting, a live butterfly exhibit at a natural history museum. Although many facets of the learning process in informal science learning institutions have been explored with populations of adults and older children, limited methodologies exist for investigating the learning of young children in these settings. More specifically, most traditional paper-and-pencil assessments commonly used in formal K-12 educational settings are not suitable for this study's target population or informal learning environment. Most young children cannot read or write, and the types of understandings they may construct from their experiences visiting informal learning settings may vary.

The first part of this chapter details the research questions, concept map assessment tool, study setting, and study sample. The second part elucidates the mixed methods approach I used to collect data. The last part of this chapter describes the quantitative and qualitative data analysis methods I used to document the learning of young children at the exhibit. This final section is divided into two parts: the first part addresses issues associated with determining the reliability and validity of the concept mapping assessment tool used in this study, and the second part addresses how young children's knowledge of butterflies was examined and evaluated.

Research Questions

As discussed in the previous section, this study aimed to both validate the use of a concept map assessment tool and also to use this tool to begin to understand what young children learn from their experiences in a live butterfly exhibit. Accordingly, the first two research questions targeted the validation of the concept map assessment tool while the remaining two questions focused on the actual use of the tool.

1. Which of three scoring systems can be used the most reliably by raters to evaluate the concept maps of young visitors to a live butterfly exhibit?
2. What is the validity of using concept maps to assess the butterfly-related knowledge of young visitors to a live butterfly exhibit? To answer this question, the following two sub-questions were posed. These questions focus on the feasibility and appropriateness of using concept maps to assess young children's butterfly-related knowledge in this informal setting.
    a) To what extent can young children construct and verbally explain propositions in their concept maps?
    b) To what extent do young children interpret the concepts in their maps as intended?
3. How does young children's butterfly-related knowledge change on family visits to a live butterfly exhibit as measured by the concept mapping assessment tool?
4. How do the concept maps of young children with and without recent experience with the live butterfly exhibit compare in terms of accuracy and content as measured by the concept mapping assessment?

Concept Maps as Alternative Assessment Tools

The multiple challenges of assessing the understanding of young children in informal settings such as science museums necessitated the development of an alternative assessment tool for use in this study. Concept maps were selected as the preferred assessment tool for this study because they have previously been used successfully to document and examine the thinking of older children in formal and informal settings (Anderson et al., 2000; James & Bixler, 2008; Klein et al., 2002; Novak

& Gowin, 1984). Furthermore, concept maps are open-ended, which allows children to communicate a range of understanding. This flexibility is an essential feature of any assessment strategy used in informal settings as visitors may construct widely different understandings from viewing the same exhibit due to varying cognitive abilities, personal interests, and prior knowledge and experiences (Allen, 2008)

Development of the Concept Mapping Assessment Tool

Concept maps are a combination of “a task that invites students to provide evidence bearing on their knowledge structure in a domain, a format for the students’ responses, and a scoring system by which students’ concept maps can be evaluated accurately and consistently” (Ruiz-Primo & Shavelson, 1996, p.573). The concept mapping task type used in this study consisted of a construct-a-map activity with concepts provided (see Appendices A and B). This particular task type was selected because past research has demonstrated its utility in demonstrating older children’s partial understandings and misconceptions (Yin et al., 2005). Furthermore, this task type allowed children to represent their understanding of butterflies without structural constraints. Although some young children may be able to create maps with a clear hierarchy, the ability to do so was not the focus of this research.

As Gallenstein (2003) suggested, instead of using written labels, the young children participating in this study constructed their concept maps using a set of eight pictures representing butterfly-related concepts (see Appendix C). The selection of these concept pictures is described in the following section. Children arranged the pictures provided to them on a magnetic white board and used a white board marker to draw linking lines between pictures as they saw fit. The use of the magnetic white board and strips of magnetic tape on the back of the pictures allowed the children to move the pictures around freely. Given that the children participating in this study did not possess well-developed writing skills, they were asked to verbally describe the relationships they identified between pictures instead of writing linking words.

Each child constructed his/her concept map and provided oral interpretations of relationships between concepts in an approximately 10-minute long "think-aloud" session. These sessions were video-taped and the completed concepts maps were photographed. "Think-alouds" have great potential for use as a tool for revealing the cognitive processes involved in concept mapping of young children and have already been used as a method for validating the interpretation of concept map scores (Shavelson & Ruiz-Primo, 2000; Yin et al., 2005).

Concept Pictures

The selection of the concept pictures began by asking a butterfly expert at the museum to provide list of science concepts presented in the exhibit; rank this list by importance for understanding the exhibit content (see Table 3-1); and construct an "expert" concept map using the top seven concepts he identified. This expert has advanced training in entomology with an emphasis on butterflies and moths and currently serves as the Assistant Curator of the McGuire Center for Lepidoptera and Biodiversity at the University of Florida. He also conducts his own research on the ecology and conservation of an endangered species of butterfly, the Miami Blue Butterfly (*Hemiargus thomasi bethunebakeri*).

As primary researcher, I have an undergraduate degree in Wildlife Ecology and Conservation and have completed advanced coursework in ecology and evolution. I conducted a content review of the butterfly exhibit's interpretative signage and panels and compiled my own list of key concepts prior to meeting with the museum's butterfly expert. I also rank-ordered my list by importance for understanding exhibit content and completed a concept map using my own top seven concepts.

The two concept maps the museum butterfly expert and I completed differed in their emphasis, with the expert's concept map focusing primarily on biodiversity and conservation of butterflies and my map focusing primarily on the morphology, behavior, and life history of butterflies. To resolve this disparity in the types of concepts emphasized in the two initial "expert" concept maps, I then developed the composite list of concepts shown in Table 3-1 that considered the content of the exhibit, the age and ability levels of the proposed study participants, and the appropriate grade-level National Science Education Standards (National Research Council, 1996). The relevant NSES K-4 content standards call for children to develop an understanding of the characteristics of organisms, the life cycles of organisms, and organisms and the environment.

In addition, the concept, insect, was added to the composite list after I noted that it did not originally appear on either the expert's or my list. This concept was added because it subsumes many of the morphology-related concepts from my original list. It was also included because the museum butterfly expert expressed concern that many exhibit visitors do not understand the basic classification of butterflies, that is, that butterflies are insects (J. Daniels, personal communication, February 17, 2009). A second concept, butterfly, was also added to this list as it was the actual topic of the concept maps.

I selected a set of preliminary pictures to match the seven concepts in the composite concept list. The composite concept list and pictures were reviewed by five experts in K-12 curriculum and instruction, educational research, and formal and informal science education. Upon review, these experts decided that one of the concepts, Species interaction, was too broad and needed to be divided into two concepts, Animal-animal interaction and Plant-animal interaction. This brought the total number of concepts in the final list up to eight (see Table 3-1). Additionally, these experts determined that two of the pictures were not suitable representations of the concepts, Metamorphosis and Rainforest strata, for young children. Based on their input, I selected pictures for the two new concepts, Animal-animal interaction and Plant-animal interaction, and selected alternate pictures for the two pre-existing concepts, Metamorphosis and Rainforest strata. Upon a second review by the two experts in formal and informal science education, these pictures were deemed suitable in their content and developmental appropriateness and became part of the set of eight concept pictures used in this study (see Appendix C).

The final list of concepts outlined in Table 3-1 was used to generate a potential proposition inventory (see Appendix D). This potential proposition inventory includes all 24 possible combinations of concepts (i.e., where a scientific relationship exists) as well as linking words describing the relationships between pairs of concepts. Given the ages of the children in this study, it was unlikely that they would use some of the more formal language of the linking phrases found in the potential proposition inventory. Accordingly, a list of children’s anticipated responses was developed along with the proposition inventory (see Appendix E). These responses were simpler, more informal ways of explaining the relationships between the pictures. The science education expert then reviewed the potential proposition inventory and the list of anticipated responses for content accuracy and developmental appropriateness.

Finally, five critical links were designated in the proposition inventory based on the top three concepts the butterfly expert deemed the most important for understanding the exhibit content prior to the study: Conservation, Biodiversity, and Species interaction. Recall that Species interaction was eliminated in favor of two more specific concepts, Animal-animal interaction and Plant-animal interaction. The five propositions that could be generated from these four concepts became the critical links in this study: Biodiversity/Conservation, Biodiversity/Animal-animal interaction, Biodiversity/Plant-animal interaction, Conservation/Animal-animal interaction, and Conservation/Plant-animal interaction.

Scoring Systems

To address research question 1, three different scoring systems were used to evaluate the children's concept maps. All three systems are based on McClure and Bell's (1990) relational scoring system, which emphasizes scoring the accuracy of propositions over their organization. In a relational scoring system, raters score concept maps by examining the proposition in a map and awarding points based on the accuracy of the propositions. The proposition scores are then summed. However, unlike McClure and Bell (1990) and McClure et al. (1999), no attention was paid to the use of directional arrows in this study. The reason for this modification is that young children may not be familiar with the use and meaning of a symbol such as a directional arrow. Each of three scoring systems is explained in the following sections; however, the procedures for scoring the maps are described in the data collection section.

Scoring system one

The first scoring system is represented in Figure 3-1. Potential scores for each proposition ranged from zero to three points. The raters examined each proposition in a child's map and made three separate decisions about the quality of the proposition. First, the raters decided if a scientifically accurate relationship exists between the linked pictures. This was accomplished by referring to the potential proposition inventory created in advance for this study (see Appendix D). Recall that the potential proposition inventory lists all 24 scientifically accurate propositions that could be constructed using the eight pictures provided to the children. Any linked pictures that matched those in this proposition inventory were automatically awarded one point.

The second decision assessed the scientific accuracy of children's verbal explanations for the pictures they chose to connect. Again, the raters referred to the potential proposition inventory and awarded a second point for each verbal explanation that represented a scientifically accurate understanding of the relationship between the concepts. Raters also had access to the list of children's anticipated responses created in advance for this study when they evaluated children's verbal explanations (see Appendix E).

The third decision assessed the completeness of children's verbal explanations. Children may demonstrate partial understanding in their verbal explanations when they describe a facet of the relationship between two concepts. The raters awarded a third point only when children's verbal explanations represented a full understanding of the relationship between two concepts.

To illuminate this scoring system further, consider the example proposition in Figure 3-2. The first picture represents the concept of "Plant-animal interaction" while the second picture represents the concept of “Metamorphosis.” Using the first scoring system, the proposition would receive one point because these two concepts are linked in the proposition inventory and it would receive an additional point for including a scientifically accurate explanation of the relationship between these two concepts. Caterpillars do eat plant material voraciously as they prepare for the next stage of their lives. The explanation, however, represents a partial understanding of the relationship because it does not directly reference how caterpillars grow and change into butterflies by repeatedly molting and eventually forming a chrysalis. Also, the explanation does not include any mention of how caterpillars are particular about what they eat and will only feed from specific host plants. The third point thus would not be awarded to this proposition and the total score for the proposition would be two points.

Scoring system two

The second scoring system is represented in Figure 3-3. This scoring system includes the three decisions described in scoring system one as well as a fourth decision regarding “critical links” in the concept maps. The use of critical links in concept mapping scoring has been identified as helpful in identifying additional information about the quality of children’s conceptual understanding (Klein et al., 2002). Recall that the critical links used in this study were the five propositions based on the concepts the butterfly expert deemed the most important for understanding the exhibit content prior to the study. The critical links included: Conservation/Biodiversity, Conservation/Animal-animal interaction, Conservation/Plant-animal interaction, Biodiversity/Animal-animal interaction, and Biodiversity/Plant-animal interaction. The raters awarded an additional point for any such critical link in a map that was accurately explained in order to identify children who showed greater understanding of the biodiversity and conservation focus of the live butterfly exhibit in their concept maps.

Consider the example proposition in Figure 3-2 again. Using the second scoring system, this proposition would receive a score of two points. As mentioned previously, the proposition Plant-animal interaction/Metamorphosis is scientifically accurate, and the explanation accompanying the proposition is also scientifically accurate (but incomplete). Two points would be awarded for these qualities. Because this proposition is not a critical link and the accompanying explanation is not complete, the proposition would not be awarded any additional points.

Scoring system three

Unlike the first two scoring systems, scoring system three evaluates concept maps using a holistic approach (see Table 3-2). In this system, I considered the overall quality of the propositions on a complete map and assigned a score for the entire map using the following rubric. The rubric focuses on the accuracy and completeness of propositions in the maps as well as the representation of critical links in the maps. Again, the inclusion of critical links was intended to identify children who showed greater understanding of the biodiversity and conservation focus of the live butterfly exhibit in their concept maps.

Consider the example proposition in Figure 3-2 once more. If half of the propositions from a child’s map consisted of such propositions (i.e., accurate but incomplete) and the other half of the propositions were not accurate, the map would receive a score of 1-“Beginning.” In contrast, if half of the propositions on a child’s map consisted of propositions like the example, and the other half of the propositions including a critical link were accurate and complete, then the map would receive a score of 2-“Basic.”

Study Setting

This study was conducted at the Butterfly Rainforest exhibit at the Florida Museum of Natural History in Gainesville, Florida in July and August 2009. The Florida Museum of Natural History is the state museum of Florida. The Butterfly Rainforest is the newest permanent exhibit at the museum and contains hundreds of live butterflies in a 6400 ft^2 screened enclosure designed to mimic a tropical rainforest. Unlike the other permanent exhibits at the museum, the Butterfly Rainforest requires an admission fee to view the exhibit. In 2008, visitors purchased over 87,000 tickets to the Butterfly Rainforest (see Figure 3-4). Adults were the largest group of visitors to the Butterfly Rainforest; however, children between the ages of three and twelve years were the next largest group of visitors.

A wide variety of butterfly species are present in the exhibit at any one time. Occasionally, some moth species are present in the exhibit as well. Many of the species originate from subtropical and tropical areas around the world. However, a few of the butterflies are native to Florida. In addition, the exhibit features many species of subtropical and tropical flowering plants, most of which are not native to Florida. These plants provide nectar and refuge for the butterflies, but are not suitable host plants. The lack of host plants prevents the butterflies from reproducing in the exhibit, which is a requirement of the exhibit’s federal and state permits.

Visitors walk along a 210-foot long curving path through the exhibit on self-guided tours. At the entrance of the exhibit, visitors may pick up laminated handouts displaying color photos, common names, and scientific names of some of the butterflies, moths, and plants found in the exhibit. Furthermore, nine interpretative panels are located along the path at varying intervals ranging from 15 to 75 feet apart. The panels include text and images describing information about the life history, behavior, adaptations, and habitat of butterflies and moths. Some topics addressed by the panels include butterfly and moth flight, drinking, feeding, coloration, reproduction, and habitat. The panels also feature bronze appliqués depicting butterfly and moth body parts such as wings, body scales, and antennae.

At least one staff member is present in the exhibit at all times with one to three additional staff members and volunteer docents present during busier times. Staff members offer interpretative programs during the release of newly emerged adult butterflies. These releases occur at scheduled times three times per day on weekends only. The staff and volunteers also answer questions and offer explanations to visitors throughout all hours of operation as needed.

Study Sample

Forty-two children aged five to seven years participated in this study. A purposeful sample from Alachua County, Florida was selected in June 2009 to include children who had a range of experiences with the Butterfly Rainforest exhibit. More specifically, 20 children had at least one previous experience with the exhibit in the past year and 22 children had not visited the exhibit in the past year. When selecting study participants, efforts were made to ensure that the two groups had similar gender and racial/ethnic compositions. However, this proved extremely difficult due the lack of response from families from diverse backgrounds as well as study participant attrition.

Participants were solicited using two approaches. In the first approach, I handed out fliers describing the study to families with young children at three public libraries in

Alachua County. One library is located in the urban center of Gainesville, a second is located in a suburban area of Gainesville, and the third is located in a rural area of the county. To reduce sampling bias, I approached every third family entering the library that appeared to include at least one child of the appropriate age range (five to seven years).

In the second study sample recruitment approach, I dropped off fliers to be distributed in kindergarten classes at eight schools with diverse student populations. (I obtained permission from the School Board of Alachua County’s public information officer and an administrator at P.K. Yonge Developmental Research School before visiting these schools.) These schools included P.K. Yonge Developmental Research School, J.J. Finley Elementary, Stephen Foster Elementary, Glen Springs Elementary, Idylwild Elementary, Terwilliger Elementary, Metcalfe Elementary, and Rawlings Elementary.

The two methods of soliciting study participants were carried out simultaneously until all participants were selected. For families that included more than one child aged five to seven years, I selected the child to participate by flipping a coin. This strategy minimized sampling bias in the ages and genders of children selected.

Signed permission forms from both parents and children were obtained before the children participated in the concept mapping sessions. Families were provided with two copies of the informed consent form, one to sign and return and one to keep for future reference. Signed forms were collected at the library, by mail (addressed envelopes were provided with return postage), or in person. Informed consent forms were stored in a locked filing cabinet along with basic demographic information about each child.

As compensation for their unpaid, voluntary participation, all children involved in the study received a complimentary butterfly coloring book. Although not intended as compensation, families who did not have annual passes to the live butterfly exhibit (all but two families) were provided with free admission for one parent and one child to help ensure that families with limited financial resources were able to participate. The Director of the museum, Dr. Douglas S. Jones, and the Program Director of the Center for Informal Science Education, Dr. Betty Dunckel, provided these free admissions (40 adult and 40 youth admissions).

Data Collection

General Procedures

After each family agreed to participate, I scheduled a date and time for the museum visit and data collection based on the availability of all those involved. Approximately six to eight family visits were scheduled each week during July and August 2009 until all data collection was complete. Ninety minutes was allotted for each visit. No more than 3 visits per day were scheduled with a 30-minute buffer in between each visit.

For each visit, I met the family at the entrance to the museum and escorted them to a private room with two tables and chairs. The room was located out of sight of the live butterfly exhibit. A magnetic white board was on one of the tables at the children's eye level. During the modeling and mapping/think-aloud sessions, I sat at the table near the white board with the participating child while the parents and other family members sat at another table a short distance away. Quiet activities such as coloring pages and puzzles were provided for siblings to use during the modeling and mapping/think-aloud sessions.

Once the family was in the room, I gave them a brief review of the procedures I intended to use. Next, I conducted the modeling session and the pre-visit mapping/think-aloud session with the participating child. The pre-mapping/think-aloud sessions were video-taped and the completed maps were photographed. In addition, completed maps were left intact. Immediately after the pre-visit mapping/think-aloud session, I escorted the family to the live butterfly exhibit. While the family took a self-guided tour the exhibit, I waited for them at the exhibit exit. No time limit was imposed on their tour of the exhibit. Immediately following their tour, I escorted the family back to the room and I conducted the post-visit mapping/think-aloud session with the participating child. Again, these sessions were video-taped and the completed maps photographed. When the child finished this last session, I thanked the child and his/her family and gave the child the butterfly coloring book.

Modeling Sessions

Creating a pictorial concept map and participating in a think-aloud session were likely new tasks for the children participating in this study, so I modeled these processes for study participants before they were asked to complete their own butterfly-related pre-visit concept maps. The topic of the example concept map was an everyday object, a sandwich. Pictures used represented eight concepts related to sandwiches (see Appendix F). I introduced the pictures to each child, and I arranged them on a large piece of white paper. As I drew the linking lines between concepts in my map, I verbally described the relationships between my linked concepts aloud such as the knife is used to take the peanut out of the jar and spread it on the bread (see Appendix G). The modeling process took approximately five minutes to complete. My completed sample "sandwich" map remained available for the children to refer to as an example when they created their own butterfly-related concept maps.

Pre-visit Mapping/Think-aloud Sessions

Immediately after the modeling session, I read the pre-visit mapping/think-aloud task to each child (see Appendix A). During the task, I presented the children with the set of eight pictures (see Appendix C). Each time the set of pictures was used, the pictures were shuffled to ensure that they were presented in random order. The color pictures were laminated and titled. In addition, the pictures had pieces of magnetic tape affixed to the back which allowed the children to arrange and re-arrange the pictures on the white board. I also provided the children with a white board marker and eraser which allowed them to draw and redraw linking lines between pairs of pictures on the white board. All sessions took approximately 10 minutes to complete.

Post-visit Mapping/Think-aloud Sessions

Immediately following the self-guided tour with their families, the children completed the post-visit mapping/think-aloud task using the concept map they created during the pre-visit mapping/think-aloud session as a starting point. They were given a white board marker and eraser and invited to modify their map to reflect any new understanding they may have developed from their experiences in the live butterfly exhibit. The post-visit mapping/ think aloud sessions also took approximately 10 minutes to complete.

Scoring Procedures

To allow raters to easily record their scores for each map using the three different scoring systems, I recorded all of the propositions for each map in the order they were created by the child on a scoring sheet created for this study (see Appendix H). In each scoring sheet, I listed names of pictures paired to form concept links (e.g., Butterfly-Metamorphosis). I then transcribed the conversation segment between the child and myself about each concept link. Instructions for the preparation and use of this scoring sheet are provided in Appendix I, and examples of a scoring sheet before and after scoring are provided in Appendices J and K. The scoring sheets were identified by numbers and did not include any information about the children.

Rater characteristics

Guided by the three scoring systems, the three raters independently generated three separate scores for each child's pre- and post-visit map using the information in the scoring sheets. I was one of the raters and the two other raters were doctoral students in science education. At the time of the study, all three raters had completed college-level coursework in ecology and entomology and had experience teaching elementary science methods courses.

I trained the other two raters to score the maps by explaining the three scoring systems and giving instructions for completing the scoring sheet (see Appendix I). I then provided the raters with the scoring sheets for the pre-and post-visit concept maps for one child selected at random (Child 2) and we scored these maps independently. When we debriefed and shared our scoring sheets, it became apparent that the other two raters were not focused enough on the accuracy of the child's explanations as they related to the targeted concepts. In order to receive the second point in both scoring systems one and two, children had to provide accurate explanations about the relationships between the concepts the linked pictures in their maps were intended to represent, not just any topic about which they might want to share. After discussing this issue, I provided the raters with the scoring sheets for the pre-and post-visit concept maps for a second child selected at random (Child 9). We again debriefed after scoring these maps and this time there were no scoring procedures that needed to be revisited. Depending on the size of the map, each map required between about five and 15 minutes to score.

Data Analysis

Part 1: Reliability and Validity of Concept Map Scores

Research question 1: Which of three scoring systems can be used the most reliably by raters to evaluate the concept maps of young visitors to a live butterfly exhibit?

As described earlier in this chapter, each map was judged by three independent raters using the three different scoring systems. The three sets of scores for each rater were entered into the SPSS data package. The scores differed in the types of data they represented: scores generated using scoring systems one and two were interval data while the scores generated using scoring system three were ordinal data. Due to the two different data types of scores, two different methods of evaluating inter-rater reliability were used. For each scoring system, the scores of each rater for all of the maps (84 total) were compared to the scores of the other two raters. For scoring systems one and two, inter-rater reliability was evaluated by computing Pearson product-moment correlation coefficients between the scores of the three raters. For scoring system three, inter-rater reliability was evaluated by computing percent agreement between the scores of the three raters.

Although understanding the reliability of concept map scores over different occasions and tasks would also be beneficial, the nature of the participants, the setting and the assessment task precluded examining test-retest reliability or internal consistency. Asking young children to complete two or more concept maps at a sitting would have most likely exceeded their cognitive capacity and patience. Similarly, there was already an existing validity threat associated with the concept mapping study. Given that the initial concept maps were completed at the museum right before participants visited the Butterfly Rainforest, exposure to the photos provided may have helped to cue these young learners and make them more aware of specific topics to focus on when they toured the exhibit. Multiple exposures to concept mapping activities focusing on these same topics and pictures could exacerbate this threat even more. Finally, the assessment task itself was not highly directed and due to its open-ended structure, a great variety of propositions were developed by study participants. The diversity of propositions generated would not allow for an analysis of internal consistency as there was not one single universal set of propositions represented across the children's maps.

Research question 2: What is the validity of using concept maps to assess the butterfly-related knowledge of young visitors to a live butterfly exhibit?

Sub-question A) To what extent can young children construct and verbally explain propositions in their concept maps? As response process evidence for validity, I reviewed my memos from the mapping sessions, videotapes of the mapping sessions, and the map scoring sheets (proposition information and concept map pictures only, not rater scores) to determine whether the children in this study were able to construct and verbally explain propositions in their concept maps.

More specifically, I looked for behaviors noted in my memos or observable from the videotapes indicating that the children were able to complete the mapping tasks such as children placing the pictures on the white board purposefully (e.g., placing certain pictures together), drawing linking lines between certain pictures, pointing to linking lines, and explaining why they drew them. Conversely, behaviors indicating difficulty completing the tasks, such as placing the pictures on the white board randomly or drawing linking lines without being able to explain them, indicated that children did not understand the task or were not able to successfully complete the mapping tasks. An abundance of “less successful” behaviors would indicate a need to modify the pre-task modeling procedure or the actual mapping tasks.

Finally, I reviewed Columns 1 and 3 of the scoring sheets for all of the children’s pre-and post-visit maps to identify the number of propositions (Column 1) they each generated during each mapping session. When reviewing these scoring sheets, I also identified and tallied instances when the children did not offer any explanations for their propositions (Column 3).

Sub-question B.) To what extent do young children interpret the concepts in their maps as intended? As response process evidence for validity, I reviewed videotapes of children’s mapping sessions and Column 3 in the scoring sheets for all of the children’s pre- and post-visit maps to compare the concepts the pictures were originally intended to represent with what children actually perceived based on their verbal explanations and descriptions of pictures during their mapping sessions. More specifically, I identified and tallied instances when children referred to the concept pictures in their verbal explanations in ways other than intended. In cases where a discrepancy was noted between the original intended and child’s actual interpretation, I also looked for patterns in how individual children interpreted the pictures and connected them to the other concept pictures in their maps.

Part 2: Children's Butterfly-related Knowledge

Research question 3: How does young children's butterfly-related knowledge change on family visits to a live butterfly exhibit as measured by the concept mapping assessment tool?

Based on the average of the total scores assigned to each map by each of the three raters using each of the three scoring system, descriptive statistics were calculated for the entire sample. More specifically, means and standard deviations were calculated for the pre-visit and post-visit map scores for each of the three scoring systems. For each scoring system, mean pre- and post-visit scores were compared using a Split-Plot Analysis of Variance (SPANOVA) to identify significant changes in butterfly-related knowledge ($p \leq 0.05$).

Furthermore, I inductively derived categories to sort the propositions children generated in their pre- and post-visit maps. In this iterative process, I reviewed the propositions developed, identified initial coding categories describing the topics of the propositions, and, when necessary, made modifications to the coding categories as new categories emerged from analysis of subsequent maps during the coding of the propositions (Bogdan & Biklen, 1998). This iterative process occurred until a "saturation" level was reached (i.e., all propositions were clearly sorted into named categories). Using pre-determined categories was not possible in this study because the assessment task is purposefully open-ended and the research examining young children's understanding of life science concepts is too limited to inform the creation of a priori categories. Finally, frequency tables for the propositions in pre- and post-visit maps of the entire sample were generated to illustrate the relative frequency of accurate conceptions and misconceptions for each major category of propositions.

Research question 4: How do the concept maps of children with and without recent experience with the live butterfly exhibit compare in terms of content and accuracy as measured by the concept mapping assessment tool?

Based on the average of the total scores assigned to each map by each of the three raters using each of the three scoring systems, descriptive statistics were calculated for the two separate groups of participants in the study: children who had visited the butterfly exhibit within the past year (n=20), and children who had not visited the butterfly exhibit in the past year (n=22). More specifically, means and standard deviations were calculated for the pre-visit and post-visit map scores for each of the three scoring systems for these two groups. These two groups were compared using the SPANOVA mentioned previously to determine whether their mean pre-visit and post-visit map scores were significantly different from each other ($p \leq 0.05$). Finally, frequency tables of the propositions in the pre-visit and post-visit maps were generated separately to compare the relative frequency of accurate conceptions and misconceptions generated by the two participant groups for each major category of propositions.

In the following chapter, I report the results of this study in two parts. The first part of the chapter details findings related to the reliability and validity of the map scores obtained using the concept map assessment. The second part of the chapter details findings related to changes in the children’s butterfly-related knowledge. This part also presents evidence of differences in children’s butterfly-related knowledge based on their recent prior experience with the live butterfly exhibit.

Table 3-1. Concepts considered for the concept mapping tasks

Butterfly expert's concept list (Rankings in parentheses)	Researcher's concept list (Rankings in parentheses)	Composite concept list	Final concept list
Behavior	Antennae	Biodiversity	Animal-animal interaction
Biodiversity (1)	Coloration (7)	Butterfly	Biodiversity
Color	Complete metamorphosis (3)	Conservation	Butterfly
Conservation (2)	Compound eyes	Insect	Conservation
Ecology (6)	Drinking	Metamorphosis	Insect
Life history	Diurnal behavior	Pollination	Metamorphosis
Metamorphosis (7)	Feeding (5)	Rainforest strata	Rainforest strata
Morphology	Flight	Species interaction	Plant-animal interaction
Pollination (4)	High diversity (1)		
Science or Scientific method (5)	Pheromones		
Species	Pollination (2)		
Species interaction (3)	Proboscis		
Strata	Puddling		
Tools	Rainforest layers (6)		
	Reproduction		
	Scales		
	Sense receptors		
	Shape		
	Temperature		
	Three part body (4)		
	Wings		

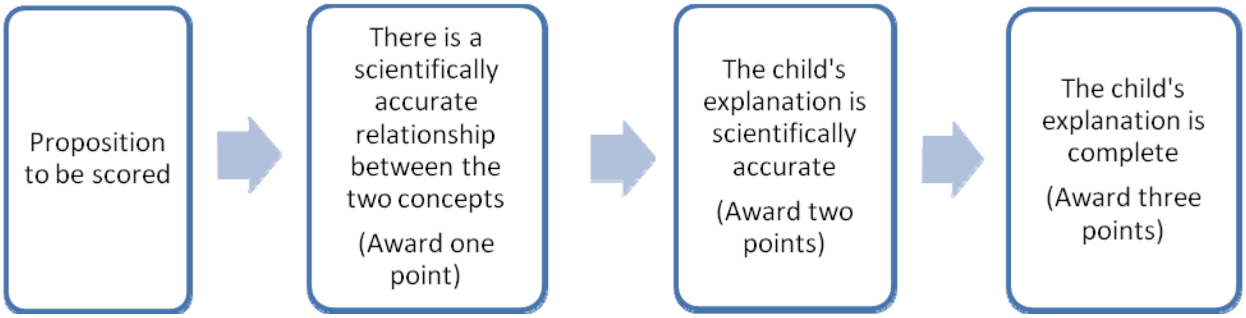
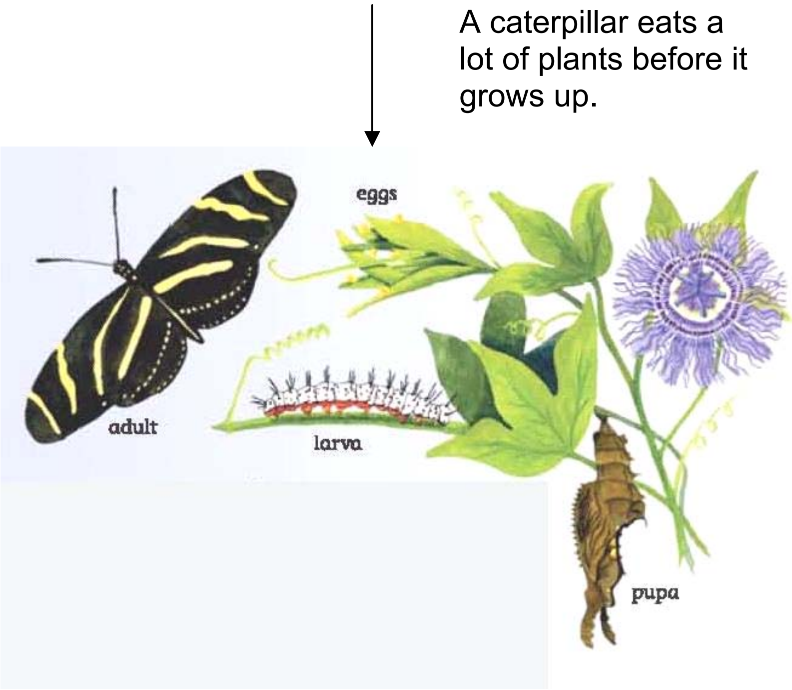


Figure 3-1. Scoring system one

Concept: Plant-animal Interaction (Title: A caterpillar eating a leaf)

A caterpillar eats a
lot of plants before it
grows up.



Concept: Metamorphosis (Title: The life of a butterfly)

Figure 3-2. Example proposition

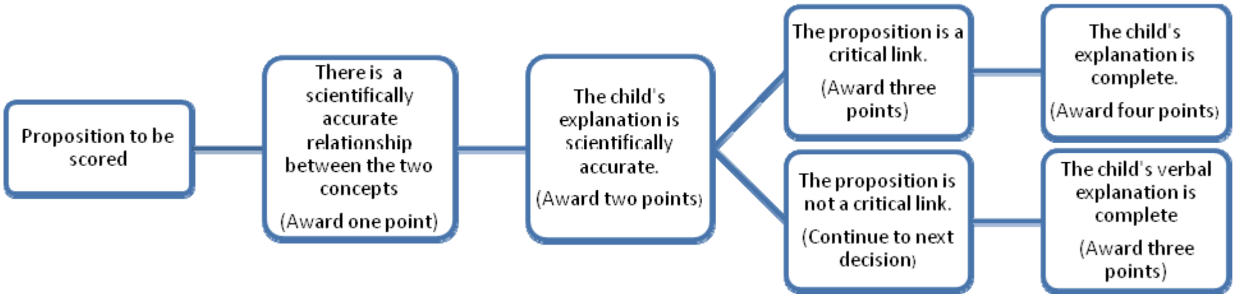


Figure 3-3. Scoring system two

Table 3-2. Holistic rubric for scoring system three

Score point value	Scoring criteria
4-Advanced	At least 80% of the propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. All three critical links are represented accurately.
3-Proficient	At least 60% of the propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. At least two of the three critical links are represented accurately.
2-Basic	At least 40% of the propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. At least one of the three critical links is represented accurately.
1-Beginning	Less than 40% propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. None of the critical links are represented accurately.

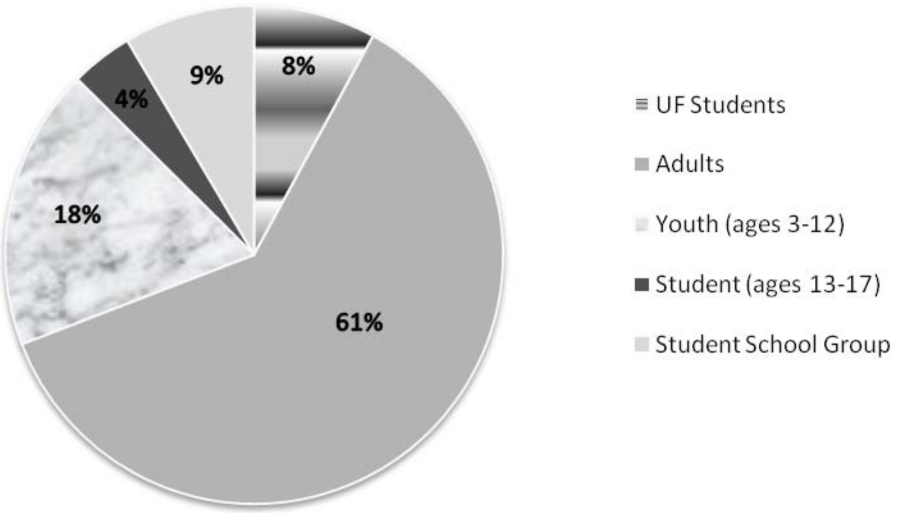


Figure 3-4. Admission tickets purchased at the Butterfly Rainforest in 2008 (K. Gerard, personal communication, January 29, 2009)

CHAPTER 4
RESULTS

Introduction

This chapter reports results of pre-and post-visit concept mapping sessions with 42 children, aged five (n=14), six (n=15), and seven years (n=13), conducted during July and August 2009 at the Florida Museum of Natural History (FLMNH). These children visited the museum's live butterfly exhibit on unguided tours with their parents, and in some cases, siblings.

The chapter first revisits the study's research questions before describing the study sample and detailing findings related to each research question. Given that the study served two purposes, the research questions and related findings are divided into two parts. Recall that the first purpose of the study was to validate the use of a novel concept mapping assessment tool with young children. In other words, the study was a pilot test of the concept mapping assessment tool with young children. The related research questions for Part 1 include:

1. Which of three scoring systems can be used the most reliably by raters to evaluate the concept maps of young visitors to a live butterfly exhibit?
2. What is the validity of using concept maps to assess the butterfly-related knowledge of young visitors to a live butterfly exhibit? Two sub-questions targeting how the children created their maps were posed to answer this question.
    a) To what extent can young children construct and verbally explain propositions in their concept maps?
    b) To what extent do young children interpret the concepts in their maps as intended?

The second purpose of this study was to use the concept mapping assessment tool to document and compare young children's butterfly-related knowledge on family visits to the museum's live butterfly exhibit. More specifically, the study compared the butterfly-related knowledge of young visitors with and without recent experience at the live butterfly exhibit. The related research questions for Part 2 include:

3. How does young children's butterfly-related knowledge change on family visits to a live butterfly exhibit as measured by the concept mapping assessment tool?
4. How do the concept maps of young children with and without recent experience with the live butterfly exhibit compare in terms of content and accuracy as measured by the concept mapping assessment tool?

Study Sample

The study sample consisted of two groups of young children recruited using the sampling methods described in Chapter 3 (see Table 4-1). The first group of children (n=20) had visited the museum's live butterfly exhibit at least once in the year prior to the study. The second group of children (n=22) had not visited the exhibit in the year prior to the study. However, these children may have visited the exhibit at some other point in their past. The parents of eight children in the second group reported bringing their children to the exhibit when they were infants or toddlers. The parents of the remaining 14 children in group 2 reported that visiting the live butterfly exhibit as part of this study was a new experience for their children.

Group 1

Study participants having recent experience with the live butterfly exhibit included 11 males and nine females (see Table 4-1). Thirty-five percent of these children belonged to a minority ethnic group. Of the males, four children were Hispanic and seven were White. Of the females, one child was African American, two were Hispanic, and six were White.

Group 2

Study participants without recent experience at the live butterfly exhibit included six males and 16 females (see Table 4-1). Twenty-seven percent of these children belonged to a minority ethnic group. Of the males, one child was Asian and five were White. Of the females, two children were African American, three were Hispanic, and 11 were White.

Child 22

Child 22 was a White male belonging to Group 1. Although five other children took down their pre-visit maps during the post-visit mapping sessions, he was the only one who failed to make a new map to replace the map he took down. Given that there was nothing to score, the raters were not able to award any points for his post-visit map. After my initial quantitative analysis of the data, I decided to remove Child 22 from the data set as an outlier.

The next sections detail findings related to the research questions for Parts 1 and 2. Quantitative and qualitative data from the children's concept mapping sessions were used to address research questions for both parts. Results for Parts 1 and 2 are then summarized separately.

Part 1: Reliability and Validity of Concept Map Scores

Research Question 1: Which of three scoring systems can be used the most reliably by raters to evaluate the concept maps of young children visiting a live butterfly exhibit?

Recall that in this pilot test of the concept mapping assessment tool, three raters independently evaluated the 42 participants' pre-and post-visit concept maps using the three different scoring systems described in Chapter 3. The next two sections review how the concept maps were scored and present results of the inter-rater reliability analyses of the map scores assigned by the three raters for each scoring system.

Scoring systems

As described in Chapter 3 and illustrated in the scoring sheet and instructions (see Appendices H and I), in both scoring systems one and two, one point was awarded for each pair of linked pictures in a concept map representing one of the 24 scientifically accurate propositions identified in the potential proposition inventory. A second point was then awarded for each accurate verbal explanation accompanying a pair of linked pictures. A final point was awarded for each complete verbal explanation accompanying a pair of linked concepts. The only difference between scoring systems one and two was that in scoring system two, an additional point was awarded for propositions representing one of five critical links. Five critical links were identified in the potential proposition inventory, and they included the following propositions: Biodiversity/conservation, biodiversity/plant-animal interactions, biodiversity/animal-animal interactions, conservation/plant-animal interactions, and conservation/animal-animal interactions. The use of critical links in the scoring system two was intended to identify children who showed greater understanding of the biodiversity and conservation focus of the live butterfly exhibit in their concept maps.

As a holistic scoring system, scoring system three required raters to consider the overall accuracy and completeness of children’s verbal explanations of propositions in their concept maps instead of making individual decisions. The presence or absence of critical links in concept maps was an essential part of this scoring system as well and was also meant to identify children who showed greater understanding of the biodiversity and conservation focus of the live butterfly exhibit in their concept maps.

However, raters used a rubric with four levels of performance (i.e., Advanced, Proficient, Basic, and Beginning) instead of using a tally system. In using this scoring system, raters judged the quality of entire concept maps (propositions and verbal explanations) rather than making individual judgments about each proposition in concept maps.

Inter-rater reliability results

Pearson product-moment correlation coefficients were generated to compare the inter-rater reliability of the scores assigned to all maps for scoring systems one and two (see Appendices L and M; Table 4-2). Given that linearity is an assumption of Pearson product-moment correlation coefficients, scatter plots were first created to determine whether the relationships between the sets of scores were linear (see Figures 4-1, 4-2, 4-3, 4-4, 4-5 and 4-6). Linearity is the constant relationship between two variables and may be positive or negative (Shavelson, 1996). In this case, a positive linear relationship was desired.

Scatter plots of the map scores for scoring systems one and two confirmed that the data set was positive and linear, thus allowing the analysis of inter-rater reliability using Pearson product-moment correlation coefficients. For this study, correlation coefficients ranging from 0.80 to 1.0 were considered "high," 0.60 to 0.79 were considered "moderate," and less than 0.60 was considered "low."

As summarized in Table 4-2, for scoring system one the correlations between raters' scores were all high, ranging from 0.94 to 0.98. Such high values indicate strong positive relationships and high levels of consistency between the scores assigned by the three raters. For scoring system two, the correlations between raters' scores were slightly lower but still considered high, ranging from 0.89 to 0.97. Again, these high values suggest strong positive relationships and consistency between the scores assigned by the three raters. The strong positive and linear relationships between the scores assigned by the three raters using scoring systems one and two were clear from the scatter plots as well. These results indicate that with minimal training raters can use scoring systems one and two to evaluate concept maps developed by young children reliably.

Scoring system three posed a dilemma. As previously mentioned, scoring system three was designed as a rubric with four discrete levels of performance (Advanced-4, Proficient-3, Basic-2, and Beginning-1). It turns out that the scores assigned using scoring system three by the raters were limited in range. Raters one and three assigned a Basic or Beginning level of performance to nearly all of the maps while rater two assigned a Basic level of performance to all 84 maps (see Appendix N).

As summarized in Table 4-3, percent agreement between the scores of the raters using scoring system three was moderate (60-79%) to high (>80%). Percent agreement between the scores of raters one and two was 78%. Percent agreement between the scores of raters one and three was 93%. Percent agreement between the scores of raters two and three was 78%. These results suggest that scoring system three was not used as consistently as scoring systems one and two.

However, further analysis using scoring system three was not continued because this scoring system poorly distinguished between different children's concept maps. In retrospect, scoring system three should have been revised to make certain that the scoring criteria more clearly differentiated between children's performance on the mapping task and perhaps a wider range of more “sensitive” scoring options should have been developed.

Research Question 2: What is the validity of using concept maps to assess the butterfly-related knowledge of young children visiting a live butterfly exhibit?

Concept maps are not a new assessment tool in either classroom or museum studies; however, they have not been widely used with young children. Understanding how young children create their concept maps thus becomes important for understanding the validity of using concept maps to assess their butterfly-related knowledge in this study. At the start of the study, it was unknown whether most young children could successfully complete the concept mapping task. This concern led to the first sub-question: To what extent can young children construct and verbally explain propositions in their concept maps? This is the first of two sub-questions posed to address the validity of using concept maps in the study. In addition, it was unknown whether most young children would interpret the eight concept pictures provided to them as intended. Thus, a second sub-question was posed to address this concern: To what extent do young children interpret the concepts in their maps as intended?

Sub-question A) To what extent can young children construct and verbally explain propositions in their concept maps?: All of the 42 children participating in the study were able to generate and give some kind of verbal explanation for at least one proposition (pair of pictures) during the pre-visit mapping sessions. In fact, the fewest number of propositions on a child’s pre-visit concept map was two and the greatest number of propositions on a pre-visit concept map was 15. The mean number of propositions generated on the pre-visit concept maps for the entire sample was 5.23 (s=0.39).

In comparison, the mean number of propositions generated on the post-visit concept maps was similar (M=5.64, s=0.42). Thirty-three of the children (78%) added to their explanations for existing propositions or created and explained new propositions during their post-visit mapping sessions. Five of the children (Child 10, 19, 28, 38, and 41), four girls and one boy, took down all of their pre-visit maps in order to create entirely new maps during the post-visit mapping sessions. Three children (Child 5, 6, and 21), two boys and one girl, examined their existing maps but decided to not share anything new during the post-visit mapping sessions. Finally, one boy (Child 22) took down his entire pre-visit map in order to draw a large, ornate butterfly during the post-mapping session. (As mentioned earlier, this child's scores were eliminated from the data set as an outlier.)

Verbal explanations did accompany more than 90% of propositions generated by the children in both their pre- and post-visit concept maps. However, given that I prompted children to provide explanations when none were spontaneously provided, this study did not generate data regarding how often young children provide unsolicited verbal explanations for propositions. Despite this limitation, there were still a few instances when no explanations were provided by children even when I prompted them. In the pre-mapping sessions, only two of the 294 total propositions created (<1%) were not accompanied by a verbal explanation. These propositions were created by two different children (Child 22 and 27), a boy and a girl. During the post-mapping sessions, a greater number of propositions were not accompanied by verbal explanations. Out of 156 total propositions generated, eight propositions (5%) lacked explanations. Four different children (Child 14, 24, 26, and 32), two girls and two boys, created these propositions.

These results indicate that the young children participating in this study were able to successfully complete the concept mapping task presented to them. Indeed, all children were able to construct at least two propositions and most constructed multiple propositions in their concept maps. Furthermore, children were able to provide verbal explanations for the vast majority of the propositions they created. These results indicate that asking young children to create concept maps using pictures and verbal explanations is a viable assessment tool for assessing their pre- and post-visit understanding in informal settings such as the live butterfly exhibit.

Sub-question B)To what extent do young children interpret the concepts in their maps as intended?: When collecting data, some differences were observed regarding how children interpreted the concept pictures provided when constructing their maps. These differences in interpretations then impacted the scoring of their maps. In 74% of the 54 concept maps containing the Plant-animal interaction concept picture, children indicated in their verbal explanations that they interpreted the picture as "butterfly larva," one of the four stages of the butterfly life cycle, rather than using it to illustrate the relationship between an animal and a plant such as that between the butterfly larva and its host plant or a butterfly adult and a nectar plant. In such cases, children often used the picture to discuss the butterfly life cycle and linked the picture to the adult Butterfly, Biodiversity, and/or Metamorphosis concept pictures. For example, in her pre-visit mapping session, Child 4 stated, "A caterpillar is one of the stages of a butterfly" when asked why she simultaneously linked the Butterfly and Plant-animal interaction concept pictures and Plant-animal interaction and Metamorphosis concept pictures, thus making two different propositions. Although her statement about the life cycle of a butterfly was correct, this explanation could not be scored as accurate for either proposition in scoring systems one or two because it did not directly relate to the intended proposition regarding the relationship between a plant and an animal.

Similarly, in 74% of the 53 concept maps containing the Conservation concept picture, children indicated in their verbal explanations that they interpreted the Conservation concept picture as a generic “butterfly habitat” rather than a butterfly habitat created by people (i.e., an example of conservation). In other words, they viewed the Conservation concept picture as strictly an example of a place butterflies could live rather than an example of an action people could take to help butterflies survive. As an indicator of their interpretation of this picture as a “place,” six children pointed out the sign stating “Butterfly Crossing” in their explanations regarding this particular picture. This makes sense because signs label important places in a child’s world (e.g., McDonald’s, their school, their street). Also, I may have encouraged this interpretation of “place” in some of my prompts when I referred to the picture as a “place.”

Consider the following conversation excerpt about the proposition linking Butterfly and Conservation during Child 6’s pre-visit mapping session as an example. The child used her observations of the sign as justification for connecting the two pictures. Also, when I questioned her about the picture, I asked how the place meets the needs of the butterfly as a habitat. The child’s explanation that the habitat provides food and shelter was accurate but could not be scored as such in scoring systems one or two because the explanation did not contain any information about the intended proposition linking the actions of people creating a habitat and helping butterflies survive.

Child 6 pre-visit map (see Figure 4.7):

Interviewer: Why do you think those two go together?

Child 6: Because the sign says “Butterfly Crossing” and there are butterflies on it.

Interviewer: Now, this place, what do you think is special about this place?

Child 6: It has plants for butterflies.

Interviewer: It has plants for butterflies. Now, the plants, what do the plants give the butterflies.

Child 6: It gives them a place to lay their eggs.

Interviewer: Very good, anything else?

Child 6: And their nectar.

Next, in 82% of the 55 concept maps containing the Rainforest strata concept picture, children indicated in their verbal explanations that they interpreted the picture holistically as a rainforest rather than the more specific intended concept of layers of a rainforest. This picture was used in combination with the Conservation picture in 13 maps as another example of a butterfly habitat. Only eight of the children discussed anything about the layers represented in the concept picture without prompting. I prompted the children to discuss the layers of the rainforest when the pace of the mapping sessions allowed.

Compare the following conversation excerpts about the proposition linking Plant-animal interaction and Rainforest strata during Child 8’s pre-visit mapping session and the proposition linking Butterfly and Rainforest strata during Child 10’s post-visit mapping sessions. In the first excerpt, Child 8 recognized that the Plant-animal interaction picture involved more than just the caterpillar but did not say anything about how the plants in the different layers of the rainforest may be host plants to different species of butterfly larvae. Also, I did not have time to interject questions about the different layers of the rainforest as I did in the second excerpt. Child 8 moved on to construct a new proposition. Thus, if Child 8 knew anything about the feeding niches of caterpillars and butterflies, it was not evident from the explanation he provided and so his explanation could not be scored as accurate in scoring systems one and two. In contrast, I was able to ask further questions in the second except in which Child 10 discussed the proposition Butterfly/Rainforest strata. In her explanation, the child discussed how different species of butterflies may live in different layers in the rainforest. As a result, this explanation was scored as accurate in scoring systems one and two.

Child 8 pre-visit concept map (see Figure 4-8):

Child 8: These things live there.

Interviewer: What are these things?

Child 8: The caterpillar

Child 8: The caterpillars eat leaves from the jungle.

Child 10 post-visit map (see Figure 4-9):

Child 10: The rainforest is like a home to lots of animals and butterflies.

Interviewer: Um hmm, that’s very true.

Interviewer: Now, do the same butterflies live in all parts of the rainforest?

Child 10: I think that they live in all parts of the rainforest.

Interviewer: Do the same kinds of butterflies live in the top, middle, and bottom of the rainforest?

Child 10: No, I think that different types of butterflies live in the canopy and other parts of the rainforest.

These results suggest that using pictures to represent specific individual concepts was problematic and unfortunately lessened the overall validity of concept map scores generated in this study. The three pictures used to represent the concepts of Plant-animal interaction, Conservation, and Rainforest strata were not interpreted as intended by about three-quarters of the children. This negatively impacted the scoring of the maps of these children because they did not receive credit for accurate understandings about butterflies when they “misinterpreted” the three pictures. This was the case for both scoring systems one and two.

In addition, scoring system two failed to clearly identify any children with a greater understanding of the biodiversity and conservation focus of the live butterfly exhibit in their concept maps because of children’s diverse interpretations of concept pictures in four of the five critical links (i.e., Conservation/biodiversity, Conservation/plant-animal interaction, Conservation/animal-animal interaction, and Biodiversity/plant-animal interaction). Concept map scores generated using scoring systems one and two were often very similar and thus the use scoring system two proved to be superfluous.

Therefore, these results suggest the need to more carefully consider the multiple possible interpretations children may have when selecting pictures to represent concepts for an open-ended concept mapping assessment tool, and to perhaps re-design the scoring systems to recognize multiple interpretations. In addition, these results suggest that the first two scoring systems might be more effective if their criteria were revised to allow more “credit” for alternate, yet correct propositions and interpretations of concept pictures.

Summary for Part 1: Reliability and Validity of Concept Map Scores

The previous sections presented evidence regarding the reliability and validity of using concept mapping to evaluate the butterfly-related knowledge of young children. The reliability evidence is based on correlations of all of the scores assigned by three raters using the first two scoring systems. These correlation results indicate that the two discrete scoring systems developed to evaluate the concept maps of children were used reliably by the raters in this study. Second, the validity evidence comes from an analysis of the responses of the children to the concept mapping task. Most children generated multiple propositions in their concept maps. Furthermore, children provided verbal explanations for nearly all of the propositions in their maps. This suggests that the children participating in this study understood, and were capable of completing, the concept mapping task as designed. However, the variability in how some of the children interpreted three of the pictures casts some doubt on the validity of the scores generated. The scoring systems may need to be re-designed to acknowledge the different ways children might interpret the concepts presented in the pictures and make provisions for recognizing links between pictures that are accurate, even if they were not anticipated.

Part 2: Children’s Butterfly-related Knowledge

The following sections summarize findings regarding the butterfly-related knowledge of the young children participating in this study. Descriptive and inferential statistics are reported for the children’s concept map scores in order to answer the remaining research questions. Additional evidence from the children’s verbal explanations for the propositions in their maps was also used to answer these research questions. More specifically, the children’s verbal explanations were analyzed using inductive category coding and the content and accuracy of six different categories of explanations identified were determined.

Research Question 3: How does the butterfly-related knowledge of young children change on family visits to a live butterfly exhibit?

To answer this question, the pre- and post-visit scores on children's concepts maps were determined and means and standard deviations were computed for pre-visit and post-visit scores using scoring systems one and two. Next, a Split Plot Analysis of Variance (SPANOVA) was performed for each scoring system (one and two) using the pre- and post-visit concept map scores. Finally, the children's verbal explanations for the propositions in their pre- and post-visit maps were sorted into six major coding categories.

Descriptive statistics

For scoring system one, the mean score on the pre-visit concept maps was 7.44 (s=3.67) and the mean score on the post-visit concept maps was 8.03 (s=3.86). After removing Child 22 as an outlier, the mean score for the pre-visit concept maps fell to 7.33 (s=3.64) and the mean score for the post-visit concept maps rose to 8.23 (s=3.69) (see Table 4-4). The range of scores using scoring system one was 2 to 18 on the pre-visit concept maps and the range of scores was 3 to 19 on the post-visit concept maps. The possible range of scores for scoring system one was 0 to 48.

For scoring system two, the mean score on the pre-visit concept map was 8.34 (s=4.18) and the mean score on the post-visit concept maps was 9.15 (s=4.46). After removing Child 22 as an outlier, the mean score for the pre-visit concept maps fell to 8.33 (s=4.17) and the mean score for the post-visit concept maps rose to 9.38 (s=4.23) (see Table 4-4). The range of scores was 2 to 20 on the pre-visit concept maps and 3 to

21 on the post-visit concept maps. The possible range of scores for scoring system two was 0 to 77.

Inferential statistics

Two 2 x 2 split-plot analyses of variance (SPANOVA) were used to analyze the map scores for the two scoring systems (see Tables 4-5 and 4-6). For both analyses, the between-subjects factor was recent experience (recent experience or no recent experience) and the within-subjects factor was mapping occasion (pre- or post-visit)..

Both SPANOVAs revealed a significant main effect for mapping occasion (scoring system one, $F=7.71$ (1, 1, 39), $p<0.01$; scoring system two, $F=9.38$ (1, 1, 39), $p<0.01$). These results indicate that there were significant differences in the pre- and post-visit map scores for both scoring systems one and two. In other words, the post-visit map scores were significantly higher than the pre-visit map scores assigned using both scoring systems one and two. These pre-post visit differences in concept map scores indicate that visiting the live butterfly exhibit on an unguided, free-choice tour with a parent does significantly improve the overall butterfly-related knowledge of young children. However, in order to determine which specific areas of butterfly-related knowledge were enhanced by these visits, further qualitative analyses were required as described in the next section.

Inductive category coding

Inductive category coding was used to sort the verbal explanations given for the propositions generated in the pre-visit and post-visit maps for all study participants (Bogdan & Biklen, 1998). After data collection, 36 codes were initially identified to categorize children’s verbal explanations for the propositions generated in the pre-visit and post-visit maps (see Appendix O). The initial 36 codes were then reviewed and condensed into six broader categories which accounted for similarities noted between some of the initial 36 codes. This round of coding did not note the accuracy of the content of children's verbal explanations.

A second round of coding was then performed focusing on the scientific accuracy of each explanation and thus accurate and inaccurate statements were identified. Results of these analyses are discussed later in this chapter. The six broader coding categories identified were: a) needs of butterflies, b) life cycle of butterflies, c) ecology of butterflies and other insects, d) diversity and classification of butterflies and other insects, e) threats to and conservation of butterflies, and f) social and maternal behavior of butterflies. Descriptions and examples of each of the six major coding categories are provided in the following sections.

A. Needs of butterflies: This coding category includes verbal explanations that referred to survival needs of butterflies. During mapping sessions, children talked about many needs of butterflies including food, air, water, and shelter. Food was the most common need discussed by children and they highlighted plants as the main sources of food for butterflies and caterpillars. In addition, plants were recognized by children as places on which butterflies can rest or lay their eggs. For example, in her post-visit mapping session, Child 19 identified pollen and nectar as food sources for adult butterflies. She also stated that butterflies use plants as "rest places." In another example, in his pre-mapping session, Child 13 stated that "the egg starts to be laid on the leaf parts of the jungle and the caterpillar crawls on the leaf parts too so it can find lots of leaf to eat so it can grow stronger to make silk to build its chrysalis."

B. Life cycle of butterflies: This coding category includes verbal explanations detailing the four major life stages of butterflies: egg, larva, pupa, and adult, or in the children's words, egg, caterpillar, chrysalis, and butterfly. Sixteen children began constructing their maps by talking about the life cycle of butterflies. For example, Child 4 began her map by stating "A caterpillar is one of the stages of a butterfly." In another example, Child 10 said that "I think that the caterpillar [picture] should go to a butterfly [picture] because butterflies start out as caterpillars." Five children even acknowledged that the lives of butterflies end. For example, Child 11 in her post-mapping session stated, "The adult butterflies make the babies but once the adult has the babies it passes on."

C. Ecology of butterflies and insects: This coding category includes verbal explanations discussing the habitats, niches, sources of predation, and defenses of butterflies and other insects. Regarding habitats, children identified rainforests and butterfly garden-type places as suitable habitats for butterflies. All but two of the children discussed at least one of these places as butterfly habitats in their pre-visit or post-visit mapping sessions. In one example, Child 3 in his post-mapping session identified the rainforest as the place where "butterflies live and caterpillars turn into butterflies." In a second example, Child 2 pointed out in her pre-visit mapping session that butterflies could live and "grow up" in the either the "wild" rainforest or the manmade butterfly habitat. Eight children acknowledged that these places were suitable habitats for other insects as well. For example, in his pre-visit mapping session Child 41 stated, "All kinds of bugs live in the rainforest." In another example, in her pre-mapping session Child

31answered, “Because bugs and butterflies live in the rainforest” when asked why she connected the Rainforest strata and Insects concept pictures.

Furthermore, 12 children discussed how different types of butterflies and insects used different parts of a habitat or fed on different plants, suggesting the idea of niches, in their pre-visit or post-visit mapping sessions. For example, Child 10 stated in her post-visit mapping session that “Some butterflies live in the top of the rainforest.” In a second example, Child 11 explained that “Some [plants] they [the butterflies] like a lot and some of them may not like that kind of nectar” in her post-visit mapping session. In a third example, Child 33 reasoned that caterpillars eat the leaves of only special kinds of plants because “some could be poisonous” in her post-mapping session. Lastly, two children, Child 13 and 33, discussed the impact of temperature differences of the various rainforest layers. For example, Child 13 stated that butterflies “put the chrysalises on the branches here [in the understory] since they are kinda high up so they don’t want their chrysalises to burn up” in his pre-visit mapping session.

Interactions between predators and prey were discussed by 35 children in their pre-visit or post-visit mapping sessions. Other insects and birds were the most commonly cited predators of butterflies. For example, Child 30 identified birds with “keen eyes” as predators of butterflies in her pre-visit mapping session. In a second example, Child 22 stated that “I know every birdie eats one kind of bug” in his pre-visit mapping session. In a third example, Child 36 stated simply that “bugs eat butterflies” in her pre-visit mapping session.

Finally, six children mentioned how prey can avoid or defend themselves against predators. During her pre-visit mapping session, Child 30 discussed how camouflage “protects you from animals that are trying to eat you.” She also explained that butterflies, moths, and other insects could have camouflage. In another example, Child 12 stated that he thought that the monarch butterfly “tastes bad” in his pre-visit mapping session. In a last example, Child 13 stated that “they [the butterflies] would hide in the chrysalis, and some animal would think that thing is poisonous and they wouldn’t eat it” in his pre-visit mapping session.

Diversity and classification of butterflies and insects: This coding category contains verbal explanations that refer to the numbers of types of butterflies and other insects, how they are classified, and their physical characteristics including how they look and how they move. Twenty-four children acknowledged that there are “a lot” of butterflies and other insects in different habitats. For example, Child 12 stated, “I know that there are lots of insects” in his pre-visit mapping session. In another example, Child 16 acknowledged that there were many different kinds of caterpillars and butterflies and then stated, “I heard the number but it’s just really big” regarding the exact number of species in her post-visit mapping session.

Furthermore, many children identified butterflies as insects or “bugs” and were able to provide some justification as to why. Their justifications usually had to do with how these organisms look (e.g., legs, antennae, etc.) or how they move (e.g., walk, crawl, fly). For example, Child 14 stated that butterflies fly and crawl and have a head, thorax, and abdomen when providing reasons why they are insects. In a second example, Child 26, during her post-visit mapping session, discussed the apparent disparity between number of legs in butterflies and caterpillars, saying “It’s weird that caterpillars have more than six legs and they turn into six legs.”

Threats to and conservation of butterflies: This coding category includes verbal explanations discussing the human and non-human threats to butterflies and how butterflies can be protected. Fifteen children discussed how butterflies needed to be protected from humans who would “squish” them or predators like birds and other insects during their pre- or post-visit mapping sessions. For example, Child 1 stated that “people have to be careful not to step on or kill butterflies” in his pre-visit mapping session. In another example, Child 41 reasoned that butterflies needed protection “in case an animal comes and tries to kill them” in his pre-visit mapping session. In a third example, Child 8 suggested that butterflies need to be protected because “bugs like to eat them” during his pre-visit mapping session. He further postulated that because of this, butterflies may go extinct.

In addition, 11 children suggested that people could help butterflies by not touching them, making a butterfly garden, keeping the predators out of the garden, and providing them with fruit. For example, Child 35 suggested simply that people “not touch them [the butterflies], not kill them [the butterflies] during his post-visit mapping interview. In a second example, Child 23 in her pre-visit mapping session reasoned that butterflies needed a butterfly garden “so that their home doesn’t get destroyed” unlike the rainforest where “people come out and chop down trees which kills butterflies’ homes and kills other butterflies.” In a third example, Child 8 reasoned in his pre-visit mapping session that the butterfly garden was “where they stop insects from doing this [eating them] to butterflies.” In a fourth example, Child 42 suggested in her post-visit mapping session that people might bring out “mangoes and bananas” to feed the butterflies in the butterfly garden.

Social and maternal behavior of butterflies: This coding category includes verbal explanations in which children compare butterflies to bees and humans in terms of social and maternal behavior. Five children discussed the need for butterflies to care for or feed “babies” in their pre-or post-visit mapping sessions. For example, Child 20 compared butterflies to bees in his pre-visit mapping session and stated, “They feed, they feed like bees, they feed the babies, they feed the themselves, they feed the queen.” In another example, Child 25 stated in her pre-visit mapping session that “they have to get nectar out for their babies and their young.”

This coding category also includes verbal explanations in which children attribute human emotions to butterflies such as happiness. According to four children, butterflies are happy when they are with their butterfly “friends” or “family,” or when they eat. For example, Child 20 stated in his pre-visit mapping session, “Butterflies are all friends. They are happy so they are friends.” In a second example, Child 28 drew a “happy” butterfly in her post-visit mapping session and said that eating made it happy.

Children generated a total of 508 verbal explanations during the pre-mapping sessions and a total of 379 verbal explanations during the post-mapping sessions. During the initial round of coding, all statements were sorted by the six major coding categories described in the preceding sections (see Tables 4-8 and 4-9). These coding categories were not overlapping and each explanation was only sorted into one category.

Determining the accuracy of children’s verbal explanations

Following the initial round of coding, I evaluated the scientific accuracy of each verbal explanation using my own knowledge of butterfly biology and ecology in a second round of coding (see Tables 4-8 and 4-9). Next, relative frequencies of accurate and inaccurate explanations for each major coding category were calculated for the total number of explanations given during both pre- and post-visit mapping occasions. This was done by dividing the number of either accurate or inaccurate explanations for a particular category by the total number of explanations given for the mapping occasion (see Tables 4-8 and 4-9). For example, children provided 508 total explanations for their propositions during the pre-visit mapping sessions. Of these, 48 explanations were determined to be accurate explanations for the category of needs of butterflies. The relative frequency for these explanations is 48/508, or 9%. (In the following sections, I report relative frequencies as percents for the sake of consistency.)

In the next sections, I first report the relative frequencies of all accurate and inaccurate explanations generated during the pre- and post-visit mapping sessions. Next, I report the relative frequencies of accurate and inaccurate explanations for the six major coding categories by mapping occasion (pre-visit or post-visit). I then make comparisons between the relative frequencies of accurate and inaccurate explanations by category for the pre- and post-visit mapping sessions. Implications of these comparisons will be discussed in the next chapter.

Accurate conceptions

Accurate explanations accounted for over 80% of all explanations generated by children during both the pre- and post-visit mapping sessions (see Tables 4-8 and 4-9). During the pre-visit mapping sessions, accurate explanations about the ecology of butterflies represented more than 30% of the total number of explanations generated (accurate or inaccurate). Accurate explanations about the life cycle of butterflies (21%), the diversity and classification of butterflies and other insects (14%), and the needs of butterflies (9%) also accounted for sizable portions of the total number of explanations.

Few accurate explanations could be attributed to the two other categories, threats and conservation of butterflies (4%) and comparing butterflies to bees and humans (2%).

During the post-visit mapping sessions, accurate explanations about the needs of butterflies and the ecology of butterflies each represented more than 20% of the total number of explanations generated. In addition, relatively large numbers of accurate explanations were generated for the following categories: diversity and classification of butterflies and other insects (17%), life cycle of butterflies (12%), and threats to and conservation of butterflies (9%). Less than one percent of accurate explanations generated in post-visit mapping sessions were related to comparing butterflies to bees and humans.

The vast majority (81%) of the content-related explanations generated by children in both the pre- and post-mapping sessions were accurate. When comparing data from the pre- and post-visit mapping sessions, the relative percentages of accurate explanations for three of the categories (the life cycle of butterflies, the needs of butterflies, and the ecology of butterflies and other insects) shifted dramatically (see Tables 4-8 and 4-9). First, accurate explanations about the life cycle of butterflies were less frequent in the post-visit mapping sessions (12% of total) than in the pre-visit mapping sessions (21% of total). Second, accurate explanations about the needs of butterflies were more frequent in the post-visit mapping sessions (21% of total) than in the pre-visit mapping sessions (9% of total). Third, accurate explanations regarding the ecology of butterflies and other insects were less frequent in the post-visit mapping sessions (21% of total) than in the pre-visit mapping sessions (31% of total). However, the relative frequency of accurate talk about the ecology category was still quite substantial in the post-visit mapping sessions.

Misconceptions

Inaccurate explanations (labeled as misconceptions in this study) represented less than 20% of the total explanations generated during both the pre- and post-visit mapping sessions (see Tables 4-8 and 4-9). During the pre-visit mapping sessions, nearly half of the inaccurate explanations generated were about the ecology of butterflies and other insects (44 out of 97 total). More specifically, these misconceptions accounted for nine percent of the total number of explanations given during the pre-mapping sessions. Misconceptions about the remaining five categories were much less frequent and together, misconceptions in all five of these categories accounted for less than 10% of the total number of explanations generated during the pre-mapping sessions.

During the post-visit mapping sessions, misconceptions about the six categories of explanations each accounted for five percent or less of the total number of explanations generated. Misconceptions about the ecology of butterflies and other insects were still the most frequent category (5% of total). Misconceptions about the needs of butterflies and the diversity and classification of butterflies and insects increased slightly between the pre- and post-visit mapping sessions (the diversity and classification of butterflies and other insects, 1% to 4% of total; the needs of butterflies, 2% to 4% of total). The relative percentages of misconceptions about the remaining three categories remained consistent between pre-and post-visit mapping sessions. Overall, misconceptions represented a relatively small percentage of the total number of explanations generated in both pre- and post-mapping sessions (19% of total).

The category of explanations containing the greatest number of misconceptions in both pre- and post-visit mapping sessions was the ecology of butterflies and other insects. However, misconceptions about this category were less frequent in the post-visit mapping sessions (5% of total) than they were in the pre-mapping sessions (9% of total). Misconceptions about the other five categories identified in the post-visit mapping sessions were either slightly more frequent (as in the case of the needs of butterflies and the diversity and classification of butterflies and other insects) or nearly the same frequency (as in the case of the life cycle of butterflies, threats and conservation of butterflies, and comparing butterflies to bees and humans).

Summary of results for research question 3

The results of the SPANOVAs and comparisons of explanations generated during the pre- and post-visit mapping sessions suggest that young children visiting live butterfly exhibits with their parents do increase their accurate understanding of butterflies. Although the relative frequency of accurate explanations generated did not appear to change pre-post, the SPANOVAs comparing the entire group's pre- and post-visit map scores demonstrate that mean post-visit map scores were significantly higher than mean pre-visit map scores. In addition, the types of accurate conceptions children shared in their post-visit mapping sessions were more heavily focused on the needs of butterflies than the life cycle of butterflies, which was their primary focus in the pre-visit mapping sessions. Furthermore, overall children generated fewer misconceptions about the ecology of butterflies and other insects during their post-visit mapping sessions than during their pre-visit mapping sessions. In summary, these results suggest that the children participating in this study did build upon and expand their conceptual knowledge of butterflies in certain key areas such as the needs of butterflies during a visit to live butterfly exhibit. Thus, one could conclude that free-choice visits to the live butterfly exhibit are effective learning experiences for young children in some key areas of conceptual understanding.

Research Question 4: How do the concept maps of children with and without recent experience with a live butterfly exhibit compare in terms of accuracy and content?

The SPANOVAs previously described did not reveal a significant main effect for recent experience nor did they identify a significant interaction between mapping occasion (pre- and post-visit) and recent experience (recent experience and no recent experience) for the scores derived using either scoring system one or two (see Tables 4-5 and 4-6). These results indicate that visits to the live butterfly exhibit are equally effective learning experiences for children with different levels of recent experience with the exhibit. In other words, children from these two experience groups had similar pre-visit map scores indicating that both groups had similar levels of butterfly-related knowledge before visiting the exhibit. These results also indicate that children from both experience groups had similar post-visit map scores and demonstrated similar levels of butterfly-related knowledge upon exiting the exhibit.

Comparisons of accuracy and content of explanations by experience group

To address research question four, additional relative frequencies (reported as percentages) of accurate and inaccurate explanations for each major coding category were calculated for the total number of explanations generated during each mapping occasion for the two different experience groups (see Tables 4-10 and 4-11). During both the pre- and post-visit mapping sessions, accurate conceptions represented 79% of the explanations generated by children with recent experience with the live butterfly exhibit (Group 1) and 82% of the explanations generated by children without recent experience with the live butterfly exhibit (Group 2). The children in both experience groups generated explanations with equivalent percentages of accurate explanations. When examining the relative percentages of accurate and inaccurate explanations by category for the two experience groups in pre- and post-visit mapping sessions, the levels of the accuracy of explanations differed for some categories between the pre- and post-visit mapping sessions. These differences are discussed in the following sections, and implications of these differences are discussed in the next chapter.

Pre-visit group differences

During the pre-visit mapping sessions, explanations generated by children from both experience groups focused heavily on the life cycle of butterflies and ecology of butterflies and other insects in their explanations (see Table 4-10). However, children in Group 2 (no recent experience) had higher relative percentages of accurate explanations related to both of these categories when compared to explanations generated by children in Group 1 (recent experience). More specifically, 22% and 32% of Group 2's total explanations represented accurate explanations of the life cycle of butterflies and the ecology of butterflies and other insects, respectively. In contrast, 16% and 27% of Group 1's total explanations represented accurate explanations of the life cycle of butterflies and the ecology of butterflies and other insects, respectively. In addition, children in Group 2 generated shared relatively fewer inaccurate explanations about the ecology of butterflies and other insects (7% of total) when compared to Group 1 (11% of total).

Additional differences between experience groups were noted in the percentages of accurate explanations identified during the pre-visit sessions for two categories: needs of butterflies and the diversity and classification of butterflies and other insects.

First, children in Group 1 (recent experience) generated more accurate explanations about the needs of butterflies (11% of total) than did children Group 2 (8% of total). Second, children in Group 1 generated fewer accurate explanations about the diversity and classification of butterflies and other insects (12% of total) than did Group 1 (16% of total).

Post-visit group differences

During post-visit mapping sessions, the categories of accurate explanations with the highest relative percentages differed by experience group (see Table 4-11). For Group 1 (recent experience), accurate explanations regarding the needs of butterflies were the most prevalent (22% of total). In contrast, the category that had the greatest percentage of accurate explanations for Group 2 (no recent experience) was the ecology of butterflies and other insects (25% of total). Although children in Group 1 had a similar relative percentage of accurate explanations for the needs of butterflies compared to children Group 2 (22% and 21%, respectively), children in Group 1 had a lower relative percentage of accurate explanations for the ecology of butterflies and other insects than Group 2 (17% and 25%, respectively).

Relative percentages of accurate explanations also differed between the two experience groups for two other categories of explanations. First, Group 1 generated fewer accurate explanations about the diversity and classification of butterflies (14% of total) and other insects compared to Group 2 (21% of total). Second, Group 1 generated more accurate explanations about the threats and conservation of butterflies compared to Group 2 (13% and 5%, respectively).

Finally, relative percentages of inaccurate explanations differed between the two experience groups for two categories of explanations. Children in Group 1 generated more misconceptions about the ecology of butterflies and other insects (8% of total) than children in Group 1 (2% of total), which was also the case for the pre-visit mapping sessions. Second, Group 1 shared fewer misconceptions related to the social and maternal behavior of butterflies (<1% of total) than Group 2 (5% of total).

Summary of results for research question 4

The results of the SPANOVAs and overall relative frequency tabulations indicate that the conceptual understanding of children with and without recent experience with the live butterfly exhibit is equally accurate when comparing both the pre- and post-visit maps of the two experience groups. The SPANOVAs yielded no significant differences in either the mean pre-visit or post-visit map scores of children with and without recent experience, indicating that the scores of these two groups are equivalent. Furthermore, the overall relative frequencies of accurate conceptions and misconceptions generated in pre- and post-mapping sessions were similar for both experience groups. These findings provide further support for the assertion that children with and children without recent experience with the live butterfly exhibit demonstrate equally accurate understandings of butterfly-related concepts in both their pre-visit and post-visit concept maps.

However, differences in the relative frequencies of accurate and inaccurate explanations by category suggest that there are some differences in the types of knowledge held by children with and children without recent experience with the live butterfly exhibit. During pre-visit mapping sessions, children who had recent experience with the exhibit demonstrated less understanding of concepts associated with the life cycle of butterflies, the ecology of butterflies and insects, and the diversity and classification of butterflies and insects than did children who had no recent exhibit experience. Conversely, “experienced” children demonstrated greater conceptual understanding of the needs of butterflies than “inexperienced” children.

During post-visit mapping sessions, children who had recent experience with the exhibit still demonstrated less understanding of concepts associated with the ecology, diversity, and classification of butterflies and other insects than children who lacked recent experience with the exhibit. On the other hand, “experienced” children demonstrated greater understanding of the threats and conservation of butterflies than “inexperienced” children.

Summary for Part 2: Children’s Butterfly-related Knowledge

Quantitative analyses: The second part of this study focused on documenting and comparing the butterfly-related knowledge of experienced and inexperienced young visitors to the live butterfly exhibit using evidence derived from concept maps. Inferential statistical analyses of the map scores assigned using the two different scoring systems (one and two) indicate that overall, post-visit map scores were significantly higher than pre-visit map scores using both scoring systems. These analyses also indicate that there are no significant differences in either the pre-visit or post-visit map scores of experienced and inexperienced children (Group 1 and Group 2). As measured by mean concept map scores of the two experience groups, children entered the study with similar levels of butterfly-related knowledge regardless of their level of prior experience with the exhibit. Similarly, children exited the study with similar levels of butterfly-related knowledge indicating that lack of recent prior experience with the exhibit does not hinder the learning of young children visiting the exhibit.

Qualitative analysis: As was the case with the results of the study’s inferential statistical analyses, the results of the inductive qualitative analysis of the content and accuracy of explanations generated by children in their pre-and post-visit concept maps also suggests that the levels of conceptual understanding of both experienced and inexperienced children were comparable. However, some interesting differences were identified in the types of accurate and inaccurate explanations generated by children in the two experience groups. Recall that, for this inductive analysis, I first identified six major categories of explanations generated including: the needs of butterflies, the life cycle of butterflies, ecology of butterflies and other insects, diversity and classification of butterflies and other insects, threats and conservation of butterflies, and comparing butterflies to bees and humans. I then compared the relative frequencies of the categories of accurate and inaccurate explanations generated for each category during pre-and post-visit mapping sessions.

These comparisons suggest that the children who did not have recent experience with the live butterfly exhibit were able to generate accurate explanations about the life cycle of butterflies, ecology of butterflies and other insects, and diversity and classification of butterflies and other insects during their pre-visit mapping sessions more often than did experienced children. Results of comparisons for another category, the needs of butterflies, yielded opposite results relative to level of prior experience. Children who had recent experience with the live butterfly exhibit shared accurate conceptions about the needs of butterflies during their pre-visit mapping sessions more frequently than did "inexperienced" children.

During the post-visit mapping sessions, children without recent experience with the exhibit still shared accurate understandings about the ecology, diversity, and classification of butterflies and other insects more frequently than did "experienced" children. On the other hand, experienced children shared accurate understandings about the threats and conservation of butterflies during the post-mapping sessions more often than inexperienced children.

Summary of Overall Findings

This research study sought to explore the use of a novel concept mapping assessment tool to document the conceptual learning of young children visiting a live butterfly exhibit at a natural history museum. Three types of scoring systems were developed and used by three independent raters; however, only the first two scoring systems yielded scores that were able to fully represent the broad range of children's responses and were found to have acceptable levels of inter-rater reliability. Children had little difficulty constructing and explaining the propositions generated in their concept maps and only a small percentage of propositions generated lacked explanations. The post-mapping task of revising initial concept maps created during the pre-visit mapping session was approached as intended by almost 80% of the children participating in this study.

The primary challenge associated with the development of the mapping task was choosing appropriate pictures for use with young children that clearly represent the concepts intended to be assessed. Still, three of concept pictures used in the study showed great variability in how they were interpreted by study participants: Conservation, Plant/animal interaction, and Rainforest strata. This finding has implications for the validity of the scores generated by raters using the two scoring systems. This issue will be discussed further in the next chapter.

Mean pre-visit and post-visit concept map scores for the entire study sample calculated by raters using the two scoring systems were significantly different from each other. This suggests that young children did increase their conceptual understandings about butterflies during their visits to the live butterfly exhibit. However, no significant difference was found between the scores of children who had recently visited the exhibit and those who had not. This finding suggests that children from both experience groups entered and exited the exhibit with comparable levels of butterfly-related knowledge despite differences in their levels of recent prior experience with the exhibit.

Finally, six coding categories were inductively derived to describe and sort the explanations children generated during the mapping sessions including: the needs of butterflies, the life cycle of butterflies, the ecology of butterflies and other insects, the diversity and classification of butterflies and other insects, the threats to and conservation of butterflies, and the social and maternal behavior of butterflies. Interestingly, some unanticipated differences in the relative frequencies of accurate and inaccurate explanations generated by “experienced” and “inexperienced” children were noted during both pre- and post-visit mapping sessions for some of these categories. These differences are summarized below and will be further discussed in the next chapter.

Pre-visit findings:

- In pre-visit mapping sessions, children who had recent exhibit experience generated accurate explanations about the needs of butterflies more often than did children who had no recent exhibit experience.
- In contrast, children who had no recent exhibit experience generated accurate explanations about the life cycle of butterflies, ecology of butterflies and other insects, and diversity and classification of butterflies and other insects more often than did children who had recent exhibit experience.

Post-visit findings:

- In post-visit mapping sessions, children who had recent exhibit experience shared accurate explanations about the threats and conservation of butterflies more often than children who had no recent exhibit experience.
- In contrast, children who had no recent exhibit experience shared accurate explanations about the ecology, diversity, and classification of butterflies and other insects more often than did children who had recent experience.
- Both groups of children showed an increase in the relative frequency of accurate explanations about the needs of butterflies in the post-visit mapping sessions.

In the final chapter, I will discuss the study's major findings related to the four research questions. Findings regarding the reliability and validity of the concept map scores will be discussed as well as findings regarding the butterfly-related knowledge of participating children in general and more specifically for children with and without recent exhibit experience. Following this discussion, I will also suggest implications of these findings for practitioners and researchers in the field of informal science education.

Table 4-1. Gender, ethnicity, and experience level of study participants

Gender	Ethnicity	Group 1-Recent experience (n)	Group 2-No recent experience (n)
Male	White	7	5
	African American	0	0
	Hispanic	4	0
	Asian	0	1
	Total	11	6
Female	White	6	11
	African American	1	2
	Hispanic	2	3
	Asian	0	0
	Total	9	16
Overall	Total	20	22

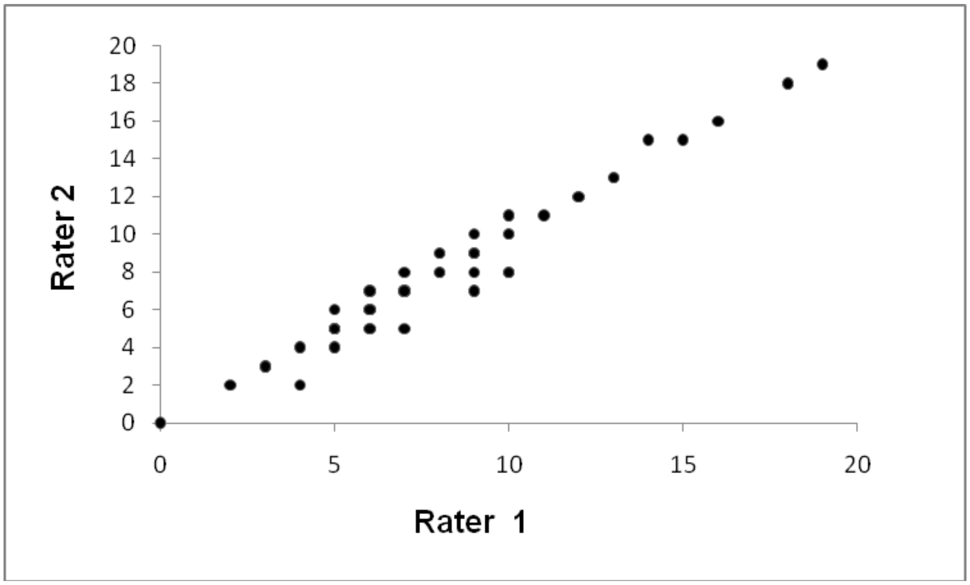


Figure 4-1. Scatter plot of map scores generated by raters 1 and 2 using scoring system one

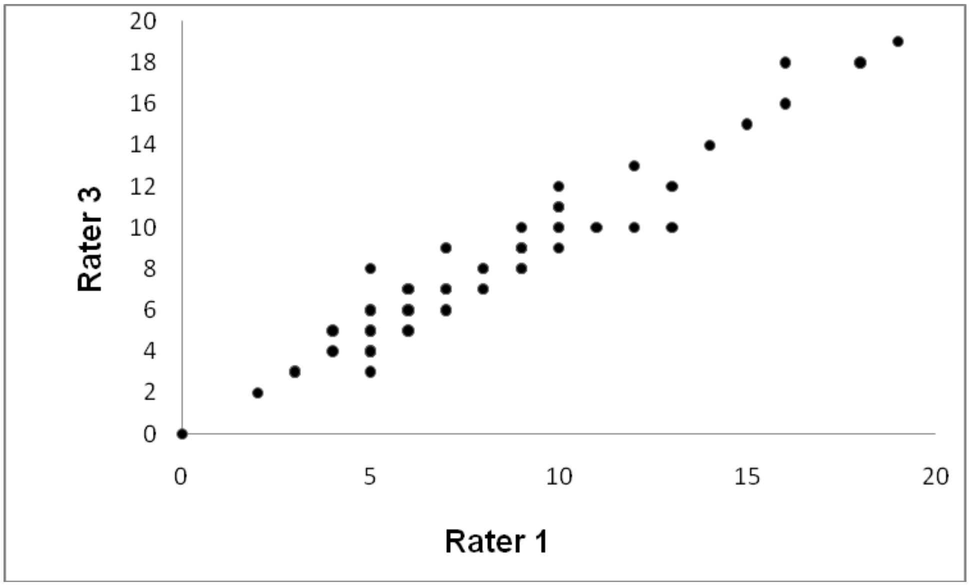


Figure 4-2. Scatter plot of map scores generated by raters 1 and 3 using scoring system one

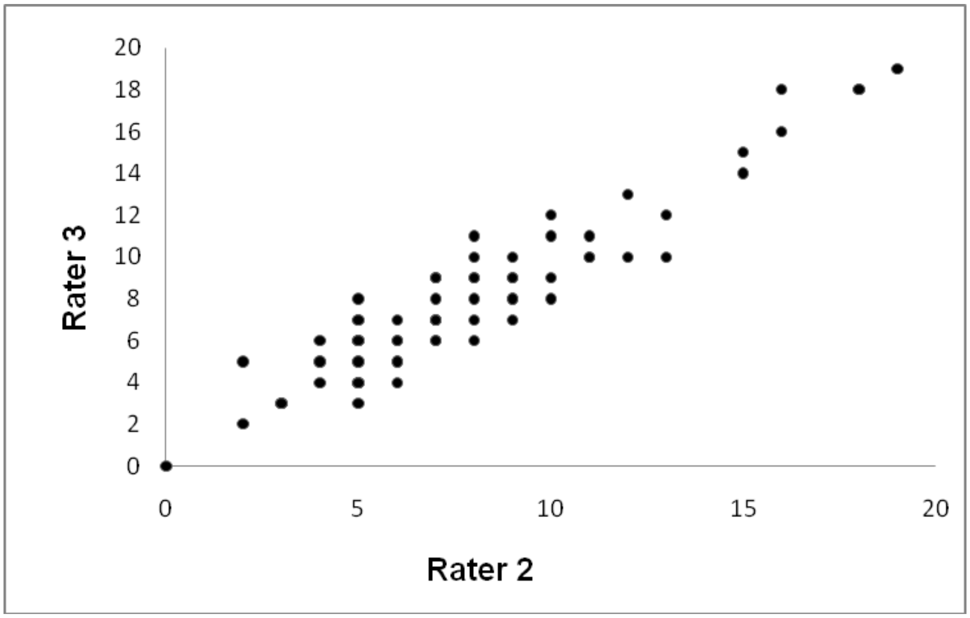


Figure 4-3. Scatter plot of map scores generated by raters 2 and 3 using scoring system one

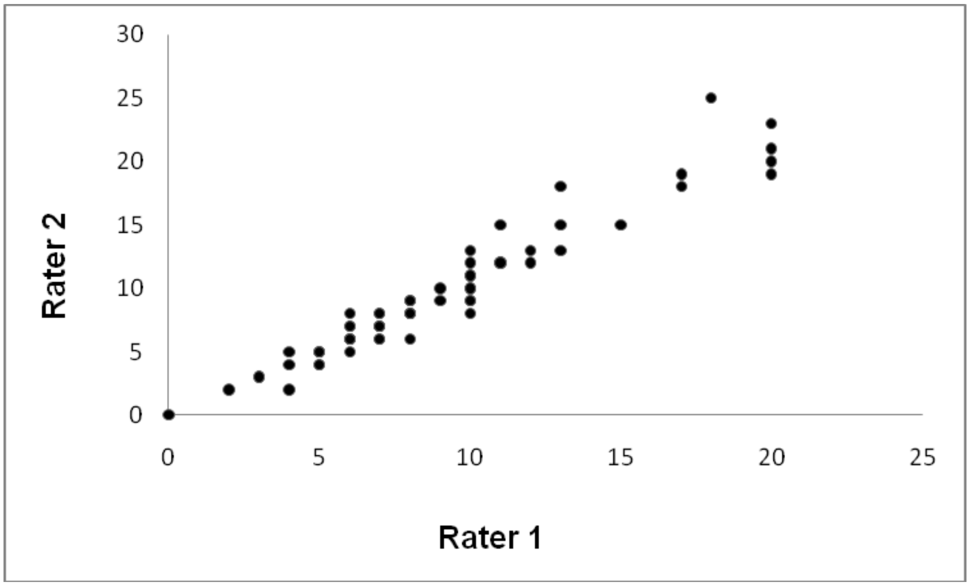


Figure 4-4. Scatter plot of map scores generated by raters 1 and 2 using scoring system two

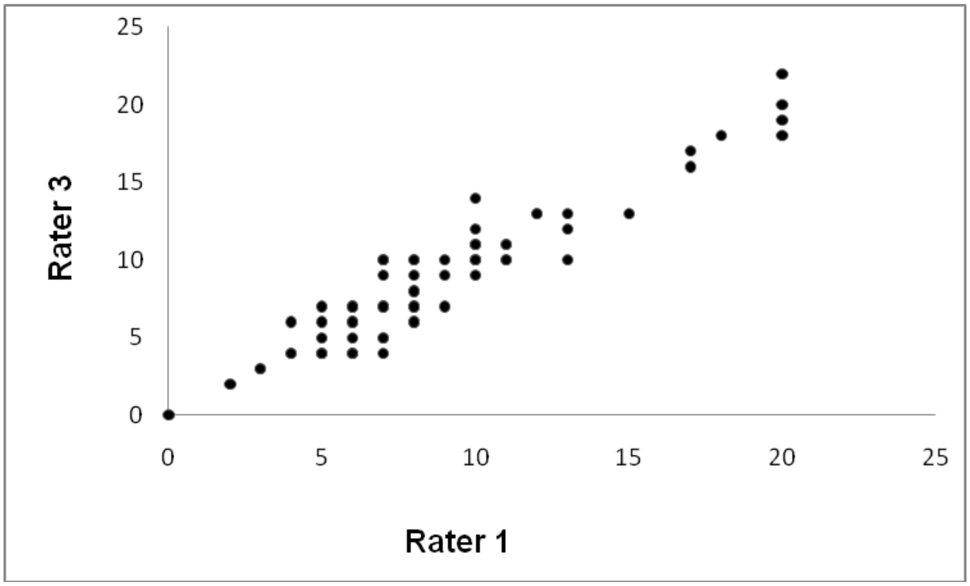


Figure 4-5. Scatter plot of map scores generated by raters 1 and 3 using scoring system two

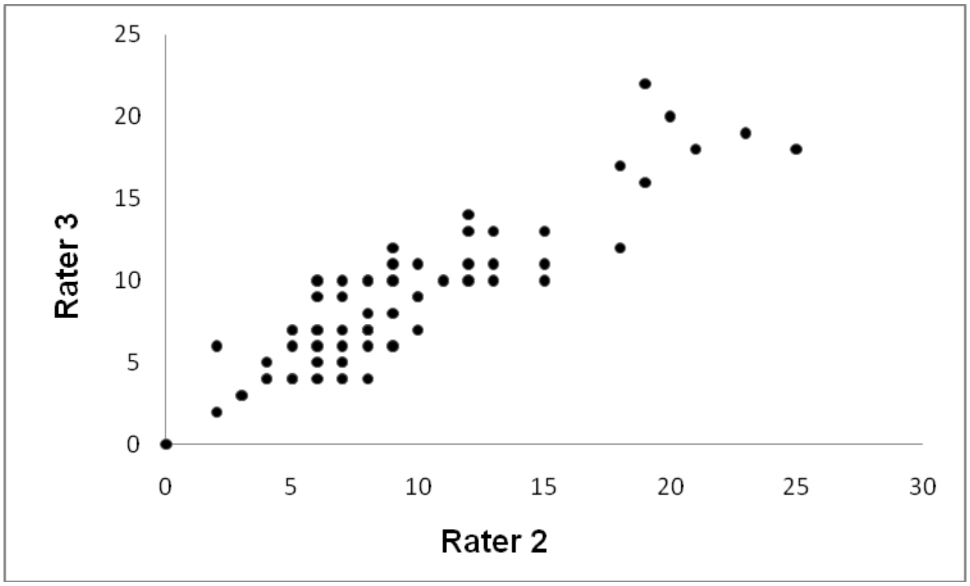


Figure 4-6. Scatter plot of map scores generated by raters 2 and 3 using scoring system two

Table 4-2. Inter-rater reliability of scores for scoring systems one and two

Scoring system	Raters	r
One	1 * 2	0.98
	1 * 3	0.97
	2 * 3	0.94
Two	1 * 2	0.97
	1 * 3	0.94
	2 * 3	0.89

Table 4-3. Inter-rater reliability of scores for scoring system three

Raters	Percent agreement
1 * 2	78%
1 * 3	93%
2 * 3	78%

Figure 4.7. Pre-visit concept map of Child 6

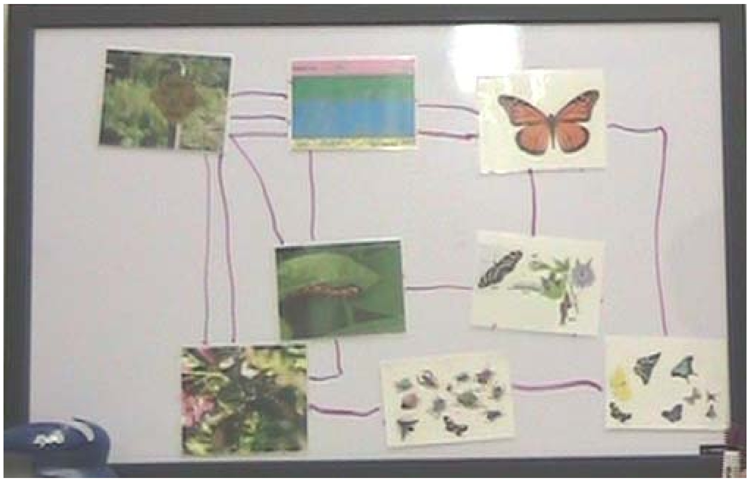


Figure 4.8. Pre-visit concept map of Child 8

Figure 4-9. Post-visit concept map of Child 10

Table 4-4. Overall means and standard deviations of map scores assigned using scoring system one and two

Map scores	n	Mean (sd)	
		Scoring System One	Scoring System Two
Pre-visit	41	7.33 (3.64)	8.33 (4.17)
Post-visit	41	8.23 (3.69)	9.38 (4.23)

Table 4-5. Test of within- and between-subject contrasts for scoring system one

Source	df	Mean Square	F	Significance
PrePost	1	20.07	7.71	<0.01
PrePost*Experience	1	4.07	1.56	0.22
Error (PrePost)	39	2.60		
Experience	1	0.59	0.02	0.88
Error	39	24.36		

Table 4-6. Test of within- and between-subject contrasts for scoring system two

Source	df	Mean Square	F	Significance
PrePost	1	30.06	9.38	<0.01
PrePost*Experience	1	1.37	0.43	0.52
Error (PrePost)	39	3.20		
Experience	1	1.84	0.06	0.81
Error	39	32.71		

Table 4-7. Total number of verbal explanations generated in pre- and post-visit mapping sessions by category (n=42)

	Number of verbal explanations	
Concept Category	Pre-visit mapping sessions	Post-visit mapping sessions
1. Needs of butterflies	60	96
2. Life cycle of butterflies	128	54
3. Ecology of butterflies and other insects	196	99
4. Diversity and classification of butterflies and other insects	78	79
5. Threats to and conservation of butterflies	24	41
6. Social and maternal behavior of butterflies	22	10
Total	508	379

Table 4-8. Relative frequencies of accurate conceptions (AC) and misconceptions (MC) identified in pre-visit mapping sessions by concept category for all participants (n=42)

Concept Category	Number of AC	Relative % of total AC+MC	Number of MC	Relative % of total AC+MC
1. Needs of butterflies	48	9%	12	2%
2. Life cycle of butterflies	108	21%	20	4%
3. Ecology of butterflies and other insects	152	31%	44	9%
4. Diversity and classification of butterflies and other insects	72	14%	6	1%
5. Threats to and conservation of butterflies	20	4%	4	<1%
6. Social and maternal behavior of butterflies	11	2%	11	2%
Total	411	81%	97	19%

Table 4-9. Relative frequencies of accurate conceptions (AC) and misconceptions (MC) for post-visit mapping sessions by category for all participants (n=42)

Concept Category	Number of AC	Relative % of total AC+MC	Number of MC	Relative % of total AC+MC.
1. Needs of butterflies	81	21%	15	4%
2. Life cycle of butterflies	44	12%	10	3%
3. Ecology of butterflies and other insects	79	21%	20	5%
4. Diversity and classification of butterflies and other insects	65	17%	14	4%
5. Threats to and conservation of butterflies	36	9%	5	1%
6. Social and maternal behavior of butterflies	1	<1%	9	2%
Total	306	81%	73	19%

Table 4-10. Relative frequency of accurate conceptions (AC) and misconceptions (MC) identified in pre-visit mapping sessions by concept category and experience level (n=42)

	Group 1-Recent experience				Group 2-No recent experience			
Concept Category	Number of AC	Relative % of total AC+MC	Number of MC	Relative % of total AC+MC	Number of AC	Relative % of total AC+MC	Number of MC	Relative % of total AC +MC
1. Needs of butterflies	27	11%	6	2%	21	8%	6	2%
2. Life cycle of butterflies	51	16%	9	4%	57	22%	11	4%
3. Ecology of butterflies and other insects	67	27%	27	11%	83	32%	19	7%
4. Diversity and classification of butterflies and other insects	30	12%	2	<1%	42	16%	4	2%
5. Threats to and conservation of butterflies	12	5%	3	1%	8	3%	1	<1%
6. Social and maternal behavior of butterflies	7	3%	6	2%	4	1%	5	2%
Total	194	79%	53	21%	215	82%	46	18%

Table 4-11. Relative frequency of accurate conceptions (AC) and misconceptions (MC) identified in post-visit mapping sessions by concept category and experience level (n=42)

	Group 1-Recent experience				Group 2-No recent experience			
Concept Category	Number of AC	Relative % of total AC+MC	Number of MC	Relative % of total AC+MC	Number of AC	Relative % of total AC+MC	Number of MC	Relative % of total AC+MC.
1. Needs of butterflies	45	22%	9	4%	36	21%	6	4%
2. Life cycle of butterflies	26	13%	5	2%	18	11%	5	3%
3. Ecology of butterflies and other insects	36	17%	17	8%	43	25%	3	2%
4. Diversity and classification of butterflies and other insects	29	14%	6	3%	36	21%	8	5%
5. Threats to and conservation of butterflies	28	13%	5	2%	8	5%	0	0%
6. Social and maternal behavior of butterflies	1	<1%	1	<1%	0	0%	8	5%
Total	165	79%	43	21%	141	82%	30	19%

CHAPTER 5
DISCUSSION AND IMPLICATIONS

Introduction

Assessing science learning in informal settings such as museums is notoriously challenging and researchers and practitioners alike have struggled with developing appropriate, valid, and reliable assessment tools for use in these settings (Bell et al., 2009). The use of traditional, quantitative multiple-choice tests has declined in recent years in favor of more authentic, qualitative assessments including interviews, observations, and concept maps. Using a wide variety of assessment tools, studies of adults and older children (older than eight years) completing short visits to informal science learning environments have yielded significant evidence that cognitive learning does occur as a result of these visits (Borun et al., 1996; Falk & Storksdieck, 2005). Furthermore, researchers have been able to document the impacts of personal (Falk et al., 1998), social (Tunnicliffe et al., 1997), and physical contexts of learning in museums on cognitive learning among adults and older children in these settings (Allen, 1997).

In contrast, the cognitive learning of younger children (younger than eight years old) in informal science learning environments has not been studied as thoroughly, due in part to the limited number of valid and reliable assessment tools available for use with young children. To date, research with this age group has largely investigated the impacts of the social context on cognitive learning, more specifically the impacts of parental scaffolding at exhibits, and has relied primarily on inferences drawn from observations to support claims of cognitive learning (Crowley, Callanan, Tenenbaum, et al., 2001).

In an effort to address the documented need for additional appropriate assessment tools for this age group, I designed and pilot-tested a concept mapping assessment task for use with young children (aged five to seven years) completing unguided, free-choice family tours of the Florida Museum of Natural History's live butterfly exhibit. Children created visual concept maps using a set of pictures provided to them and provided verbal explanations of the connections between pictures in their maps. Qualitative and quantitative data generated from the concept mapping sessions provided insights into the reliability and validity of this assessment tool and helped to document changes in the children's butterfly-related knowledge resulting from the visit.

This chapter will discuss study findings and their implications related to the four research questions framing the study. In brief, the first two research questions address aspects of the reliability and validity of the concept mapping assessment tool developed for use in this study. The third and fourth research questions focus on comparisons of the pre and post-visit butterfly-related knowledge of young children and comparisons of the pre and post-visit butterfly-related knowledge of two groups of children participating in the study (i.e., those with and without recent experience).

Research Question 1

Which of three scoring systems can be used the most reliably by raters to evaluate the concept maps of young visitors to a live butterfly exhibit?: For the purposes of this study, efficiency is defined as the amount of time and cognitive effort different scoring systems require of raters. Some scoring systems require raters to make complex decisions about many different map components, and thus, they require more time and cognitive effort. This study explored the use of three different scoring systems for evaluating the pre-visit and post-visit concept maps generated by the study's child participants. Although all three scoring systems were based on McClure and Bell's (1990) relational scoring system for concept maps which emphasizes scoring the accuracy of propositions over their organization, these three scoring systems differed significantly in their efficiency. Of the three systems used, scoring system one was found to be the most efficient.

As in past studies (Klein et al., 2002; McClure et al., 1999; Shavelson & Ruiz-Primo, 2000), the data generated in this study revealed that even though scoring systems one and two differed in their levels of efficiency, their inter-rater reliability did not differ dramatically. Scoring system two had only slightly lower inter-rater reliabilities even though it was less efficient than scoring system one. (Recall that scoring system two required raters to make two additional scoring decisions, one of which was a complex decision regarding the completeness of each verbal explanation.) The inter-rater reliabilities of scoring systems one and two were both high (>0.80) demonstrating that the three raters in this study were able to score maps consistently. These high reliability coefficients are similar to those reported in other studies in which multiple raters were used to score concept maps (Rice et al.,1998; Rye & Rubba, 2002; Schau et al., 2001; Stoddart et al., 2000; Van Ziele, Lenaerts, & Wieme, 2004).

While McClure et al. (1999) had some success developing a holistic concept map scoring system that could be used reliably by different raters, this was not the case in this study. Scoring system three (a holistic scoring system) poorly distinguished between the maps of children. Given that this current study provided insight into children's butterfly-related knowledge, in the future it may be possible to develop a holistic scoring system that is more representative of typical children's concept maps.

As is the case with other research investigating concept mapping assessment tools for use with adults or older children (Ruiz-Primo et al., 1997), I sought to balance concerns over feasibility and psychometrics. Given the young age of the participants in my study and the limited time they were able to spend on the tasks, I only administered one version of the pre- and post-visit concept mapping assessment tasks during the children's museum visit to determine inter-rater reliability. This approach precluded determining equivalent forms or test-retest reliability. In addition, I chose an open-ended task which precluded determining the internal consistency of the assessment. Given that this study contributed useful information about what young children typically know about butterflies, future work may use this information to develop a more directed task such as a fill-in-the-map task. Subsequent studies may then be able to compare the variability of scores due to task sampling.

Research Question 2

What is the validity of using concept maps to assess the butterfly-related knowledge of young visitors to a live butterfly exhibit?: Many forms of evidence may be used to support the validity claims for an assessment's scores (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). Content-based evidence for validity is one such form of evidence. Similar to the validation approaches used by other researchers (Klein et al., 2002; Shavelson & Ruiz-Primo, 2000), I consulted a content expert, an expert in butterfly ecology and conservation, in order to develop an initial list of appropriate concepts for potential use in the concept mapping assessment. His input helped ensure that the concepts represented in the pictures were aligned with the conceptual content included in the signage and displays of the live butterfly exhibit. With the young ages of my study participants in mind, I also consulted with pedagogical experts to ensure that the concepts selected were appropriate for the cognitive development levels of five-to-seven year olds.

Response process-based evidence is another form of evidence that may be used to support validity claims (AERA, APA, & NCME, 1999). I documented how children constructed their concept maps, including the number of propositions they generated by connecting pairs of pictures and whether their picture pairs were accompanied by verbal explanations to form a complete proposition. This response process-based evidence indicates that the young children in my study were capable of successfully completing the concept mapping assessment.

Furthermore, I documented the range of interpretations children generated in their verbal explanations of picture pairs. This response process evidence suggests that many children did not interpret three of the eight pictures (i.e., Conservation, Plant-animal interaction, Rainforest strata) as intended. Given this finding, it may be inferred that, for some children, the assessment did not fully reflect their complete butterfly-related knowledge. Also, the scoring systems using critical links did not clearly identify children who showed greater understanding of the biodiversity and conservation focus of the live butterfly exhibit due to the variety of interpretations children possessed related to concept pictures in four of the five critical links (i.e., Conservation/biodiversity, Conservation/plant-animal interaction, Conservation/animal-animal interaction, and Biodiversity/plant-animal interaction).

Previous researchers have not deemed it necessary to document children’s interpretations of pictures in their concept mapping assessments (Figueiredo et al., 2004; Hunter et al., 2008; & Monroe-Ossi et al., 2008). This study clearly demonstrates the need to collect data to ascertain young children's possible interpretations of pictures prior to their use in concept mapping assessments. Future work should incorporate methods such as think-aloud interviews or focus groups when selecting pictures for possible use in concept mapping assessments in order to ensure more consistent interpretation by children.

Research Question 3

How does young children's butterfly-related knowledge change on family visits to a live butterfly exhibit as measured by the concept mapping assessment tool?: Even though science museum experiences are frequently cited as sources of early informal science learning, the existing research regarding young children and science museums has provided limited direct evidence that cognitive learning occurs during visits to these informal settings. Researchers have listened to conversations and observed the behavior of young children at science museum exhibits with their families but few have actually spoken to individual children (Crowley, Callanan, Jipson, et al., 2001). In contrast to these less direct approaches, this study aimed to directly assess young children's butterfly-related learning resulting from a museum exhibit visit. Children shared their thoughts by verbalizing the relationships they saw between picture pairs as they created their concept maps.

Three major findings were identified when children's concept maps and their accompanying verbal explanations were analyzed. First, the majority of children participating in this study possessed accurate and wide-ranging knowledge about butterflies even before visiting the live butterfly exhibit regardless of prior exhibit experience. Second, most children's knowledge increased after their visit to the live butterfly exhibit regardless of prior exhibit experience. Third, children's prior knowledge and subsequent knowledge focused more heavily on some aspects of butterfly-related knowledge than others, specifically the life cycle of butterflies, the needs of butterflies, and the ecology of butterflies and other insects.

These findings support the growing recognition that young children are more competent in learning science than previously thought (Duschl, Schweingruber, & Shouse, 2007). The children in this study not only demonstrated their pre-existing knowledge of the biology and ecology of butterflies, but also their apparent ability to construct new knowledge about butterflies on unguided family visits of the live butterfly exhibit. The following sections discuss findings regarding the children's prior knowledge of butterflies and their knowledge of butterflies following their visits to the exhibit.

Finding 1: Prior Knowledge of Butterflies (Pre-visit Concept Maps)

The pre-visit concept mapping assessment task results indicated that children entered the exhibit with largely accurate knowledge about butterflies. In fact, about 80% of the verbal explanations generated by study participants were accurate. In particular, the majority of children demonstrated a clear understanding of the four-stage life cycle of butterflies before they even entered the exhibit. This finding contrasts with the results of previous research suggesting that young children have limited understanding of the life cycle of butterflies (Barrow, 2002; Shepardson, 1997). Unlike the children in this study, children in past studies ignored the egg or pupal stages of development prior to receiving formal instruction in the butterfly life cycle.

Second, most children in this study already possessed accurate knowledge about the ecology of butterflies and insects, such as habitats, niches, sources of predation, and defenses of butterflies and other insects. Nearly 80% of children's verbal explanations about ecological topics in the pre-visit concept maps were accurate. However, inaccuracies about ecological topics accounted for nearly one-tenth of all verbal explanations generated in the pre-visit concept maps and represented the largest group of misconceptions documented in this study (pre or post-visit). Although these accurate conception/misconception results are mixed, the results of this study still do not seem to fully support claims from past studies regarding the limited ecological knowledge young children possess prior to formal instruction in ecological concepts (Gallegos, 1994; Leach et al., 1996b; O'Byrne, 2008; Snaddon et al., 2008; Strommen, 1995).

Third, prior to the museum visit, more than half of the children in this study showed understanding of concepts including the diversity and classification of butterflies and other insects. Children in this study used many of the same physical characteristics (e.g., legs, wings) to classify insects that are described in other studies documenting children's insect descriptions before receiving formal instruction in the classification of insects (Barrow, 2002; Prokop et al., 2007; Shepardson, 2002; Trowbridge & Mintzes, 1998; Tunnicliffe & Reiss, 1999.

The results of this study suggest that most young children entered the live butterfly exhibit with an accurate understanding of many aspects of butterfly biology and ecology, which raises the question: Where does this knowledge originate? In the state of Florida, the growth and development of living things, including butterflies, is specifically addressed as a benchmark in the K-2 Sunshine State Standards in science content that were in place at the time of this study (Florida Department of Education, 1996). Some ecological concepts, such as habitats and food chains, are also addressed in these K-2 science standards. However, other topics with which young children in this study were already familiar, such as the classification of invertebrate animals, are not directly addressed in these grade-level standards. These findings suggest that children in this study may have been taught about the life cycle of butterflies and some ecological concepts at school. Thus, prior formal instruction may be one possible explanation for the relatively high levels of butterfly-related knowledge children possessed regarding the life cycle and ecology of butterflies and other insects.

Alternatively, “bugs” and the butterfly life cycle are both topics that are frequently represented in young children’s trade books, websites, and television shows. Thus vicarious experiences via media exposure may be another possible source of prior knowledge regarding the life cycle, diversity, and classification of butterflies. A last possible source of prior knowledge about all three of these butterfly-related concept topics about butterflies (and other insects) may be direct at-home learning experiences such as observing backyard butterflies, raising caterpillars, or tending gardens with family members. Gainesville and the surrounding local area include many garden centers and a butterfly farm that supply both the necessary butterfly host and nectar plants and caterpillars to private citizens.

Finding 2: Subsequent Knowledge of Butterflies (Post-visit Concept Maps)

Following their experiences in the live butterfly exhibit, the majority of children in this study showed a greater level of understanding of butterfly-related knowledge on the post-visit concept mapping assessment task. After their visit, most children shared more accurate knowledge about the needs of butterflies, including food, air, water, and shelter. In post mapping sessions, most children also recognized that plants provide butterflies with their main food sources and shelter. The most probable explanation for this increased post-visit knowledge is that children focused on these topics during their observations when visiting the exhibit. At any given time during the summer months of data collection for the study, scores of butterflies could be seen from any location in the exhibit feeding on nectar and fruit, and resting on plants. This explicit evidence of the needs of butterflies seemed to be readily observed by children during their visits.

These findings agree with those reported by Myers et al. (2004) that the physiological needs of a favorite organism in a similar setting (a zoo) are readily noted by young children. Myers et al. (2004) also asserted, along with other researchers, that young children rarely grasp the ecological needs of living things such as habitats and species interactions (Gallegos, 1994; Leach et al., 1996b; O'Byrne, 2008; Snaddon et al., 2008; Strommen, 1995). The findings of this study seem to suggest otherwise. In this study, most children showed a greater understanding of the ecology of butterflies and other insects (e.g., habitats, niches, and predation) following their exhibit visit. Children still frequently shared accurate understandings about the ecology of butterflies and other insects in their post-visit mapping sessions. In addition, the portion of verbal explanations representing misconceptions about the ecology of butterflies and other insects shared in the post-visit mapping sessions declined compared to pre-visit misconception frequencies. One possible explanation for this finding is that children were able to walk through and directly experience the “natural” environment of the live butterfly exhibit rather than observing it through a window or from an observation deck as is the case at many zoos. This difference may have allowed children to more closely attend to the features of the habitat and interactions between butterflies and plants throughout the exhibit.

Finding 3: Comparisons of Prior and Subsequent Knowledge

In this study, most children's prior knowledge of butterflies was concentrated in three areas: the life cycle of butterflies, the ecology of butterflies and other insects, and the classification and diversity of butterflies and other insects. As discussed previously, potential sources of this prior knowledge include formal instruction, vicarious experiences through media exposure, and direct at-home experiences with butterflies. Following their visits to the live butterfly exhibit, most children demonstrated greater subsequent knowledge of butterflies that was more heavily focused on the needs of butterflies compared to their prior knowledge. As suggested earlier, this may have been due to the many opportunities the exhibit provided for children to closely and easily observe butterflies interacting with plants, flowers, and fruits to satisfy some of their survival needs.

Research Question 4

How do the concept maps of young children with and without recent experience with the live butterfly exhibit compare in terms of accuracy and content as measured by the concept mapping assessment tool?: Previous research exploring the personal context of learning suggests that the prior experiences of visitors may promote different levels of cognitive learning at museum exhibits (Falk & Dierking, 1992; 2000). In particular, some studies of schoolchildren (older than ten years old) on field trips have documented a novelty effect in which children who receive a pre-visit orientation show greater understanding of science content after a museum visit than those who do not receive a novelty-reducing orientation (Anderson & Lucas, 1997; Anderson et al., 2000; Kubota & Olstad, 1991). To explore the impact of a possible novelty effect on the learning of young children visiting the live butterfly exhibit, I compared the post-visit butterfly-related knowledge of two groups of children between the ages of five and seven years. The first group of young children (n=20) had visited the live butterfly exhibit at least once in the year prior to the study. The second group of children (n=22) had not visited the exhibit in the year prior to the study. The results of this comparison are discussed in the following paragraphs.

Evidence against a Novelty Effect

On any particular day, the live butterfly exhibit used in this study contains hundreds of free-flying, brightly-colored tropical and subtropical butterflies. Despite this highly stimulating sensory environment, the quantitative analysis of concept map data in this study revealed no significant differences in the post-visit concept map scores of the two groups of children. Experienced and inexperienced children exited the exhibit with similar levels of butterfly-related knowledge. This finding suggests the absence of a novelty effect when it comes to butterfly-related knowledge gains resulting from the exhibit visit. This assertion is further supported by a comparison of the level of accuracy of verbal explanations generated by these two groups. During the post-visit mapping sessions, accurate explanations represented 79% of the explanations given by children with recent experience and 82% of the explanations given by children without recent experience. Thus, children showed equivalent levels of accuracy in their explanations regardless of past exhibit experience.

Evidence for a Novelty Effect

The qualitative analysis of young children’s verbal explanations generated in this study provided intriguing evidence suggesting that children do acquire different types of knowledge depending on their level of prior experience with the exhibit. During the post-visit mapping sessions, both groups (experienced vs. inexperienced) demonstrated similar levels of understanding in two concept categories: the needs of butterflies and the life cycle of butterflies. Of the two groups, children who had recent experience with the exhibit showed greater post-visit understanding in one category: the threats to and conservation of butterflies. Alternatively, children who did not have recent experience with the exhibit showed greater post-visit understanding in two categories: the ecology of butterflies and other insects and the diversity and classification of butterflies and other insects.

These conflicting findings do not provide clear support for a novelty effect; however, the novelty effect described in previous informal settings research related to differences in the amount of science content learned rather than differences in the types of science content learned from museum visits (Anderson & Lucas, 1997; Anderson et al., 2000; Kubota & Olstad, 1991). The findings of this study do seem to support the findings of Falk and Adelman (2003) who documented the positive benefits of prior knowledge and prior interest on learning during museum visits. Although the children in this study did not significantly differ in their entry levels of butterfly-related knowledge, they did differ in some of the types of butterfly-related knowledge they initially possessed. Prior to their visits, inexperienced children in this study showed greater understanding of butterfly-related knowledge regarding the ecology, diversity, and classification of butterflies and other insects than experienced children did. Furthermore, inexperienced children maintained higher levels of knowledge regarding these areas of butterfly-related knowledge in post-visit comparisons. Possible reasons for these differences vary, but one likely reason is that the “inexperienced” children were not truly inexperienced in regards to butterflies despite the fact they had not visited the live butterfly exhibit recently. Although not explored in this study, these children and their parents may have participated in other butterfly-related learning experiences in the home and community before participating in this study.

Limitations

Not unlike many museum researchers, I found the recruitment of a diverse study sample to be challenging. The bulk of my study sample consisted of White children (69%), while the demographic composition of the population of the state of Florida is only 60% White. Furthermore, most of these children resided in the suburban areas of Gainesville. Gainesville is a unique city in that it maintains and adds to its urban green space. Accordingly, many species of animals, butterflies included, inhabit the protected natural areas in and around the city including forests, creeksides, and prairies. Therefore, the concept mapping assessment task and information gained from its use reflects the responses of a fairly homogeneous group of children and are not generalizable to the larger, more ethnically and geographically diverse population of children aged five to seven years.

This study cannot fully account for the interesting differences noted in the types of butterfly-related knowledge held by the two experience groups of young children. As discussed earlier, additional factors such as varying degrees of initial interest in butterflies and prior participation in other butterfly-related activities besides recent visits to the live butterfly exhibit (e.g., classroom instruction, firsthand experience raising caterpillars, viewing backyard butterflies) may have contributed to these differences. Furthermore, families who chose to participate in this study may have had greater prior interest in butterflies than the general population. In other words, the sample of children in this study may have suffered from a degree of “self-selection” in that families with greater interest in butterflies may have more readily opted to participate.

Finally, this study did not investigate the types of learning behaviors or content of conversations of children and their parents when visiting the live butterfly exhibit. As a result, it is unknown if specific interactions between families or interactions with exhibit features (e.g., signage, live butterflies) are related to changes in children’s butterfly-related knowledge. It is also unknown if mere participation in this study noticeably changed the learning behaviors or conversations of children and their parents when visiting the exhibit. Parents may have cued children to attend to aspects of the exhibit or signage they thought related to the concepts addressed in the pre-visit concept mapping assessment task.

Implications for Practice

Using Concept Maps as an Assessment Tool

The findings of this study further support the use of concept maps to assess the science learning of young children (Figueiredo, Lopes, Firmino, & de Sousa, 2004; Fleer, 1996; Hunter, Monroe-Ossi, & Fountain, 2008). Assessment of science learning in any setting may be diagnostic, formative, or summative. First, concept maps may be used as diagnostic assessment tools to ascertain young children’s familiarity with science concepts and plan future learning experiences in classrooms or museum exhibits. Second, concept maps may be used as formative or summative assessment tools to determine the success of learning experiences in classrooms or museum exhibits in promoting science learning goals and objectives and inform modifications of such learning experiences to maximize their success.

Ideally, assessment practices are imbedded in young children’s science learning experiences. In the classroom setting, concept maps may be imbedded as a learning center at which young children create their own concept maps using pictures or objects and provide verbal explanations to a more experienced peer or adult. In the museum setting, concept maps may be embedded as an interactive exhibit feature at which young children create their own concept maps using pictures or objects and provide verbal explanations to a trained docent. Educational technology such as the software program Kidspiration paired with audio recording equipment may facilitate more formal assessment by allowing young children to create concept maps which can be reviewed later by educators in either formal or informal settings.

Planning Future Learning Experiences about Butterflies

The findings of this study add to the body of research regarding what young children know about living things in general and insects and butterflies more specifically (Barrow, 2002; Prokop et al., 2007; Shepardson, 2002; Trowbridge & Mintzes, 1988; Tunnicliffe & Reiss, 1999). The young children participating in this study demonstrated knowledge of many aspects of butterfly biology and ecology even before their visits to the live butterfly exhibit. In particular, they were highly knowledgeable about the four-stage life cycle of butterflies. The ecology of butterflies was another area in which they demonstrated significant initial understanding; however as a group, they also possessed a sizable number of misconceptions about concepts such as species interactions. Children in this study also had some initial understanding of the diversity and classification of butterflies and insects.

Insights from this study could be used to guide the development of future learning experiences about butterflies in informal and formal settings for this age group. These experiences should extend young children’s learning about butterflies by moving them beyond what they already know well (i.e., the life cycle of butterflies) to gaining greater understanding of topics they have not yet mastered (i.e., ecology, diversity, classification, threats to, and conservation of butterflies and other insects).

Documents such as the National Science Education Standards (NRC, 1996) and Benchmarks for Scientific Literacy (American Association for the Advancement of Science, 1993) emphasize that even young children should actively learn about scientific processes and concepts through “hands-on” experiences. Although the children’s visits to the live butterfly exhibit in this study were not “hands-on,” they were experiential and provided children with opportunities for direct interaction with hundreds of free-flying butterflies in a “natural” environment. These experiences seem to be ideal for promoting young children’s understanding about the survival needs of butterflies. Indeed, it would be hard to imagine anyone visiting this exhibit without noticing a butterfly feeding on fruit or nectar. Considering this fact, how might subsequent direct experiences with butterflies in children’s homes, schools, and communities help them to extend their learning regarding other areas of butterfly-related knowledge?

Home and school-based experiences

One possible strategy informal science educators could consider is the development of a menu of suggested extension learning activities for parents and teachers that focus on additional aspects of butterfly biology and ecology. Such activities could range from planting home or school butterfly gardens to participating in existing citizen science projects focused on butterflies (e.g., Journey North or Monarch Watch). If resources are available, informal science educators could also offer families and teachers the option of borrowing materials for observing and identifying butterflies at home or on school grounds such as hand-lenses, beginner's field guides, and simple field logs suitable for young children. Structured outdoor learning experiences such as directed observation and identification of living organisms have been shown to be effective tools for promoting children's learning in many fields of life science, including biology, ecology, and conservation (Dillon et al., 2006; Cronin-Jones, 2000; Carrier, 2009).

Museum-based experiences

Another possibility is the development of stronger networks that allow informal science educators to collaborate with museum professionals in order to incorporate more hands-on experiences for young children and families who visit live butterfly exhibits. For example, designers of live butterfly exhibits could incorporate more age-appropriate interactive features for young children such as puzzles, sensory experiences, and role plays. As a second example, trained docents could accompany young children on their visits to the exhibit and set up learning stations that allow children to observe preserved specimens or interact with models, toys, or stuffed animals that clearly illustrate key morphological features used in classification.

As a third example, museum interpretation staff could develop more explicit and age-appropriate signage and reusable handouts that prompt families with young children to make their own observations of and inferences about butterfly morphology and behavior. Children could receive a small incentive such as a stamp in a "passport" for completing each observation with their parents and earn a larger incentive such as a recognition certificate for completing a desired number of observations. This "passport" program may have the added benefit of encouraging repeat visits by families with young children.

The need for youth education regarding conservation of butterflies

Although few studies have examined the effectiveness of conservation education efforts targeting young children, the existing research suggests that, even at an early age, young children are in the process of developing an awareness of the negative impacts humans can have on the environment (Davis, 2009). Similarly, some children in this study exhibited a clear awareness of a number of ways people can harm butterflies directly (e.g., stepping on them) or indirectly (e.g., cutting down the rainforest). However, in this study children shared fewer ideas about how to conserve butterflies locally (e.g., planting a butterfly garden) and none were able to articulate ways they could help conserve butterflies in distant rainforests. How might young children extend their awareness of how to conserve butterflies, particularly those residing in distant rainforests, and perhaps even take personal action to conserve butterflies locally and far away?

One idea would be the inclusion of a list of suggested monthly or weekly “personal acts of kindness” towards butterflies, including those living in distant rainforests, in the menu of extension activities designed for families with young children outlined earlier. These suggested actions could focus on directly benefiting local butterfly species such as planting nectar and host plants and reducing the use of pesticides in gardens (i.e., children may help spray soapy water on pest insects). Alternatively, these acts could also focus on benefiting distant rainforest butterfly species such as heeding the 3Rs (Reduce, Reuse, Recycle) and donating a part of a child’s allowance to organizations purchasing rainforest acreage for conservation purposes, such as the Rainforest Action Network or The Nature Conservancy.

Finally, the existing research suggests that repeated, direct experiences in "natural" environments such as the live butterfly exhibit may be one avenue for improving both cognitive and affective outcomes important for the adoption of conservation behaviors (Littledyke, 2008). Although attitudes such as care and concern for butterflies were not explored in this study, the children who had recently visited the exhibit showed greater awareness of the threats to and conservation of butterflies than those who were new to the exhibit. Thus, simply visiting a live butterfly exhibit repeatedly may help raise children's awareness of the conservation needs of butterflies. Accordingly, it may be beneficial to allow families to purchase shorter, month-long passes to the live butterfly exhibit at a reduced rate during times they plan on conducting repeat visits (e.g., summer vacation).

Implications for Educational Research

Science Education Research

The process used to develop and validate the concept mapping assessment tool in this study has three important implications for science education researchers and science teachers conducting "action research" exploring the use of novel alternative assessment tools. First, the "checklist" scoring systems (one and two) more clearly captured the breadth of children's butterfly-related knowledge compared to the holistic scoring system (three) using in this study. This finding indicates that holistic scoring systems may not be as useful in assessing complex performance tasks as checklist scoring systems.

Second, scores generated using "checklist" scoring systems were equivalent in terms of their reliability and validity, indicating that increased efficiency does not sacrifice either reliability or validity. Thus, the wise use of raters' time may be a legitimate priority for researchers in developing scoring systems for new assessment tools. Similarly, science teachers may not need to select the most time-consuming method of assessing student work products in their action research, saving their own time.

Third, the lack of prior work documenting how children may interpret the pictures in the concept mapping assessment tool reduced the validity of the scores generated in this study. Researchers interested in developing visually-based alternative assessment tools should conduct pilot studies of children’s interpretations of any graphic (e.g., photograph, line drawing) used. Lists of acceptable alternate interpretations of graphics should be used in subsequent work developing scoring procedures.

Informal Science Education Research

This study documented the effectiveness of a promising new cognitive learning assessment tool suitable for use with young children in informal settings. Many researchers in the informal science education community have focused on how parents scaffold children’s learning in informal science learning environments. They have undertaken the commendable task of documenting the variety of verbal and nonverbal behaviors parents use to support their children’s learning at exhibits. The assessment tool developed and tested in this study may be useful in extending this existing research to include the documentation of differences in children’s learning resulting from such parental scaffolding. For example, Crowley, Callanan, Tenenbaum, et al. (2001) noted that parents offer more explanations to their boys than to their girls at exhibits. They posited that this bias results in girls being at a greater disadvantage for learning science from museum exhibits. With further testing, the assessment tool developed for this study could help researchers investigate and document the learning impacts of gender differences in levels of parental scaffolding.

The approach to assessment used in this study is unlikely to remain the only alternative assessment tool that “works” in informal science learning environments. Given the pressure from funding authorities to document learning gains in informal settings among visitors of all ages, researchers should continue to develop and validate other alternative assessments appropriate for the developmental level of young children such as drawings and card sorts. As mentioned previously, the difficult part of using any type of non-reading based alternative assessments in research is determining reliability and validity. As indicated in this study, one pilot test may not be enough to accomplish this task. However, the reward of efforts to expand the types of assessment tools available may be a broader understanding of the impacts of family visits to informal science learning environments on the emergent scientific literacy of young children.

Finally, the live butterfly exhibit in this study differs from the traditional and interactive exhibits described in the vast majority of informal science education studies. Traditional exhibits (e.g., natural history dioramas) are those in which visitors may only passively view objects behind glass walls or in display cases while interactive exhibits (e.g., interactive models) are those in which visitors have the option to actively touch and use objects. The live butterfly exhibit, on the other hand, is an example of an experiential, or “walk through,” exhibit in which visitors’ senses are bombarded by a wide array of sight, sounds, and smells as they walk through the exhibit. Future work could compare the relative impact of three different types of butterfly exhibit experiences (traditional, interactive, experiential) on the butterfly-related knowledge of young children. Although the aesthetics of a live butterfly exhibit are unparalleled, the higher costs of building and maintaining a live butterfly exhibit may not be warranted by the learning benefits when compared to other exhibit types.

Conservation Education Research

As shown in this study, even young children can be competent science learners in informal settings and already possess accurate understandings about certain aspects of butterfly-related knowledge. Given what is already known about the personal context of learning, it is likely that the knowledge of children matches their interest to some degree. Considering the high level of interest young children seem to have in butterflies, butterflies may be a fitting topic of focus for future conservation education research efforts targeting young children and their awareness, attitudes, and behaviors related to the conservation of local and global biodiversity.

Recommendations for Future Research

In order to gain a more accurate picture of how young children from different ethnic groups or geographic areas (e.g., urban vs. rural vs. suburban) might perform on the concept mapping assessment task and to document the impact of visiting a live butterfly exhibit on children’s butterfly-related knowledge, future studies should include additional recruitment strategies targeting ethnically and geographically diverse children. Such strategies may include visiting urban and rural schools and attending community events in neighborhoods with high minority populations as face-to-face conversations may be more persuasive recruitment tools than paper fliers. Given the importance of understanding the emergent scientific literacy of all young children in order to plan and implement more effective formal and informal science learning experiences (Bell et al., 2009; Duschl et al., 2007), studies involving more diverse samples would be highly useful.

Future studies should document young children's prior participation in everyday (e.g., reading books about butterflies, observing backyard butterflies), informal (e.g., visits to museums with live butterfly exhibits), and formal learning experiences (e.g., classroom instruction about the butterfly life cycle) regarding butterflies as well as determine their and their parents' entry-level interest in butterflies. Although self-reports have their limitations, they might be one method that could be employed. More specifically, children, parents, and teachers might provide evidence of children's prior interest and experiences related to butterflies in surveys or interviews which could then be correlated with children's prior knowledge regarding butterflies. Such work would be useful in further substantiating Falk and Adelman's (2003) research regarding the combined positive impacts of prior interest and prior knowledge on learning in museums and Crowley and colleagues' research regarding children and parents' shared areas of interest and expertise (Crowley & Jacobs, 2002; Crowley & Palmquist, 2007).

In addition, future studies could complement the pre-post visit concept mapping tasks with additional "during visit" data collection documenting the behaviors and conversations of children and parents as they walk through the live butterfly exhibit. Data regarding how children and parents explore the exhibit, including time spent in the exhibit and types of interactions with each other and exhibit features, might be useful in identifying the length of exhibit visits and specific types of interactions that maximize children's cognitive learning in the exhibit. Such data would also be useful for determining if completing the concept mapping assessment task significantly alters parent-child interactions in the exhibit.

Future studies could explore methods of modifying the successful scoring systems (one and two) and the scoring sheet used in this study. One possible modification might be to include a section in the scoring sheet to document children's initial impressions of the concept pictures prior to the mapping sessions. This modification may allow children to share numerous valid interpretations of the concept pictures and thus may provide additional evidence of their understanding of butterflies. A second possible modification might include incorporating a section in the scoring sheet to document accurate and inaccurate verbal explanations generated by children for each concept category. This modification would allow for the identification of statistically significant changes in accurate and inaccurate explanations generated in the pre-and post-visit mapping sessions overall and for each concept category. Such data would be particularly useful for understanding the impact of visiting a live butterfly exhibit on children's butterfly-related knowledge if analyzed with additional data about children's prior interest in and experiences with butterflies.

Future researchers may also wish to consider including additional mapping opportunities to further examine the reliability of this concept mapping assessment tool. For instance, children may be able to complete two concept maps at a time (perhaps with a short break in between) if fewer concept pictures (i.e., four or five) were provided to them. This would allow researchers to examine parallel forms reliability. In another example, children and families may be willing to return to the museum to complete a second post-visit mapping session a week or a month following their visit. This would allow researchers to examine test-retest reliability as well as give researchers some indication of the lasting impact of a live butterfly exhibit visit on young children's understanding of butterflies.

Finally, knowledge gains may not be the only benefits of a live butterfly exhibit visit for young children. However, there are no available methods of assessing changes in young children's awareness, attitudes, and behaviors related to conservation to date. Future researchers may wish to develop additional measures to detect affective and behavioral impacts of a live butterfly exhibit for young children. Young children may be a critical audience for conservation-focused informal settings such as the live butterfly exhibit, and butterflies may be a charismatic microfauna that can be used to help achieve broader conservation goals (i.e. reducing consumption, preserving contiguous tracts of rainforest habitat).

Conclusion

Until recently, young children were an invisible audience in informal science education research (but definitely not in informal science learning environments). They were viewed in some cases as inseparable from the family unit. The knowledge of young children is not easy to assess in formal settings and in informal settings, this problem is exacerbated by the well-documented difficulties of assessing learning of all age groups in free-choice settings such as museums.

Speaking and interacting with young children directly seems to be the most logical approach to assessment with this age group given their developing literacy skills. Although this assessment approach requires a certain degree of finesse (and humor), practitioners and researchers alike have much to gain from interacting with this group directly. Practitioners can gain insights into what types of knowledge should be extended in programming for young children and what types need not be addressed, putting their resources to best use. Researchers can become more flexible in their view of assessment in informal settings, particularly with young children. This flexibility may ultimately lead to a greater understanding of the impacts that everyday, family visits to informal settings have on the emergent scientific literacy of young children.

APPENDIX A
PRE-VISIT MAPPING AND THINK-ALOUD PROTOCOL

- Turn video camera on.
- Hold out the set of pictures for the child to see.
- Say, “Look at all of these pictures.”
- Say, “There are eight pictures in all.”
- Say, “Listen carefully while I tell you what each picture is.”
- In random order, read the titles of the pictures to the child and place each picture in front of the child on the table in a row.
- Say, “You are going to make a thinking map with them like I did with the sandwich pictures.”
- Say, “Think about how the pictures might have something to do with butterflies.”
- Pick up the butterfly picture and place it above the row of the seven remaining pictures
- Wait 10 seconds.
- Say, “Now, you are going to show me how you think the pictures might have something to do with butterflies.”
- Say, “As you show me, I want you to tell me why you think the pictures go together.”
- Say, “You do not have to use all of the pictures. Remember, there is no wrong or right way to make a thinking map because it’s your thinking map.”
- Hand the child the pictures and the white board marker, and direct him/her to the white board.
- Place the eraser on the marker tray of the white board, and say, “Here is the eraser if you need it.”
- Listen and watch the child mapping. Smile and nod encouragingly. If needed, repeat directions or show example concept map.
- If at the end, the child does not explain a proposition, point to the proposition and say, “Why did you draw a line between these two pictures.” Repeat until all propositions have been explained.
- Turn the video camera off.
- Take a picture of the concept map on the white board.

APPENDIX B
POST-VIST MAPPING AND THINK-ALOUD PROTOCOL

- Turn the video camera on.
- Say, “Think about what just you saw and did in the Butterfly Rainforest.”
- Wait 10 seconds.
- Say, “Look carefully at the thinking map you made earlier.”
- Say, “Are there any changes you think you can make to your map because of what just you saw and did?”
- Wait 10 seconds.
- Say, “Now, you are going to show me how you can change your map.”
- Say, “You can remove pictures, add pictures, move around pictures, erase the lines between pictures, or draw new lines between pictures.”
- Say, “When you do any of these things to your map, I want you to tell me why you made these changes.”
- Hand the child the white board marker and direct him/her to the white board.
- Listen and watch the child while he/she makes changes. Smile and nod encouragingly. If needed, repeat directions.
- If at the end, the child does not explain a change to his/her concept map, point to the change and say, “Why did you make this change to your thinking map?” Repeat until all changes have been explained.
- Turn the video camera off.
- Thank the child and parents for their help.
- Take a photograph of the concept map on the white board

APPENDIX C
PICTURES FOR BUTTERFLY-RELATED CONCEPTS

1) Concept: Biodiversity (Title: Picture of many kinds of butterflies)

(Source: http://office.microsoft.com/en-us/clipart/default.aspx. Last accessed May, 2009.)

2) Concept: Conservation (Title: Picture of a safe place people made for butterflies)

(Source: http://www.kidsgardening.com/growingideas/PROJECTS/feb04/pg2.html Last accessed May, 2009)

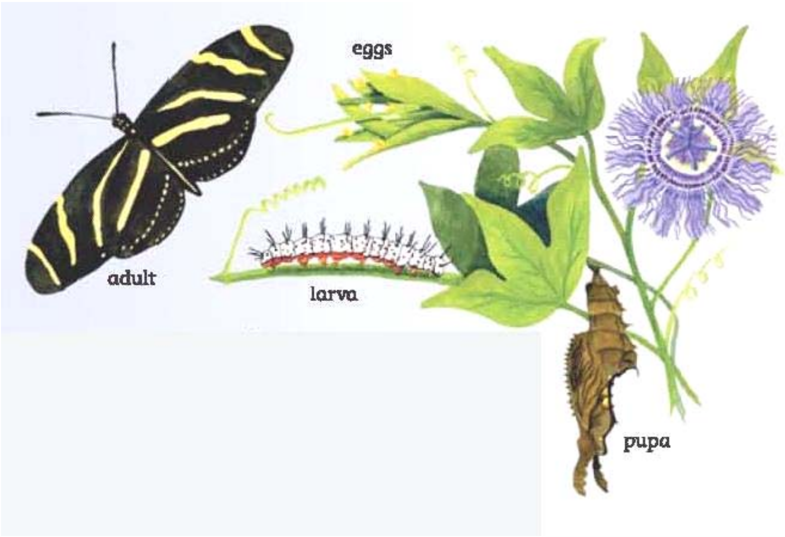
3) Concept: Plant-animal interaction (Title: Picture of a caterpillar eating a leaf)

(Source: http://office.microsoft.com/en-us/clipart/default.aspx. Last accessed, March 2009)

4) Concept: Animal-animal interaction (Title: Picture of a praying mantis eating a butterfly)

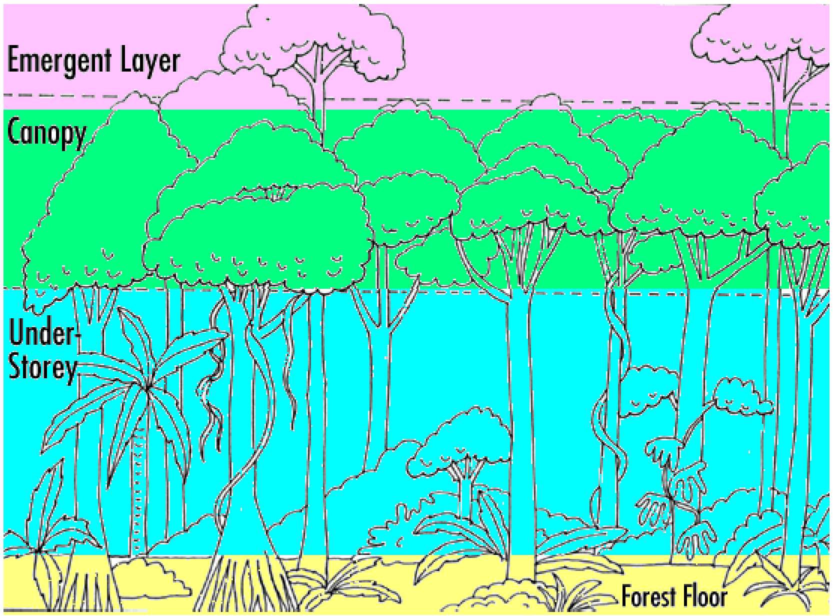
(Source: http://office.microsoft.com/en-us/clipart/default.aspx. Last accessed May, 2009.)

5) Concept: Metamorphosis (Title: Picture of the life of a butterfly)



(Source: http://www.flmnh.ufl.edu/wildflower/brochure_side1_butterflies.pdf)

6) Concept: Rainforest strata (Title: Picture of rainforest layers)



(Source:
http://nhs.needham.k12.ma.us/cur/environment/Envir00_01/p4/arspd4/layers.htm
Last accessed May, 2009)

7) Concept: Insect (Title: Pictures of an insect)

(Source: http://office.microsoft.com/en-us/clipart/default.aspx. Last accessed, May 2009)

8) Concept: Butterfly (Title: Picture of a butterfly)

(Source: http://office.microsoft.com/en-us/clipart/default.aspx. Last accessed, March 2009)

APPENDIX D
POTENTIAL PROPOSITION INVENTORY

	Biodiversity	**Conservation**	**Plant-Animal Interaction**	**Animal-Animal Interaction**	**Metamorphosis**	**Rainforest strata**	**Insect**	**Butterfly**
Biodiversity		Maintains*	Preserves*	Preserves*		Differs by, have high	Have high	Have high
Conservation	Maintains*		Should be conserved*	Should be conserved*		Should be conserved	Should be conserved	Should be conserved
Plant-Animal Interaction	Preserves*	Should be conserved *		Various responses	Different stages play a role in different interactions, (host and nectar) plants meet needs of different stages	Interactions differ by	Play a role in	Play a role in
Animal-Animal Interaction	Preserves*	Should be conserved*	Various responses		Different stages play a role in different interactions	Interactions differ by	Play a role in	Play a role in
Metamorphosis			Different stages play a role in different interactions, (host and nectar) plants meet needs of different stages	Different stages play a role in different interactions			Undergo	Undergo
Rainforest strata	Differs by	Should be conserved	Interactions differ by	Interactions differ by			Differ by	Differ by
Insect	Have high	Should be conserved	Play a role in	Play a role in	Undergo	Differ by		
Butterfly	Have high	Should be conserved	Play a role in	Play a role in	Undergo	Differ by		

*Critical Links=Biodiversity/Conservation, Biodiversity/Plant-animal interactions, Biodiversity/Animal-animal interactions, Conservation/Plant-animal interactions, Conservation/Animal-animal interactions

APPENDIX E
CHILDREN'S ANTICIPATED RESPONSES FOR POTENTIAL PROPOSITION INVENTORY

Proposition Linking Words	Anticipated Responses
Maintain	Helps Is important for
Preserves	Keeps Helps Saves
Should be conserved	Should be protected Should be saved
Differs by	Is different Is not the same
Play a role in	Is important for Helps
High	Many A lot
Meet needs of	Gives food Gives a place to live
Undergo	Go through changes in Turns into (different stages) in Becomes (different stages) in
Is a type of	Is a kind of Is a sort of

APPENDIX F
PICTURES FOR MODELING SESSIONS

Picture 1
Title: Picture of a peanut butter and jelly sandwich

(Source: http://office.microsoft.com/en-us/clipart/default.aspx. Last accessed, May 2009)

Picture 2
Title: Picture of a jar of peanut butter

(Source: http://office.microsoft.com/en-us/clipart/default.aspx. Last accessed, May 2009)

Picture 3
Title: Picture of a jar of jelly

(Source: http://office.microsoft.com/en-us/clipart/default.aspx. Last accessed, May 2009)

Picture 4
Title: Picture of a butter knife

(Source: http://office.microsoft.com/en-us/clipart/default.aspx. Last accessed, May 2009)

Picture 5
Title: Picture of a kitchen

(Source: http://office.microsoft.com/en-us/clipart/default.aspx. Last accessed, May 2009)

Picture 6
Title: Picture of a girl eating a sandwich

(Source: http://office.microsoft.com/en-us/clipart/default.aspx. Last accessed, May 2009)

Picture 7
Title: Picture of a lunch

(Source: http://office.microsoft.com/en-us/clipart/default.aspx. Last accessed, May 2009)

Picture 8
Title: Picture of toasted slices of bread

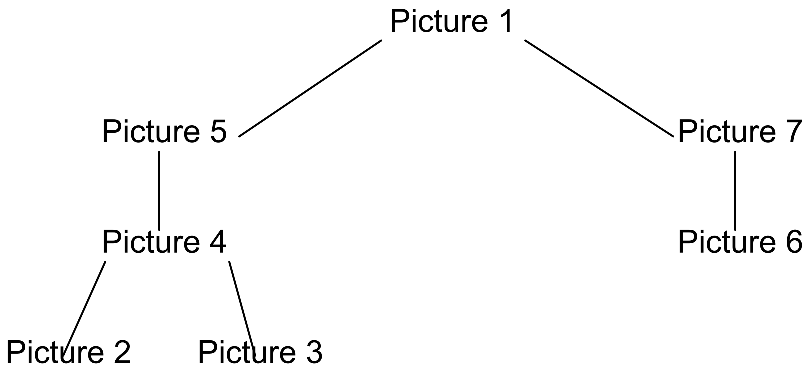
(Source: http://office.microsoft.com/en-us/clipart/default.aspx. Last accessed, May 2009)

APPENDIX G
PROTOCOL FOR MODELING SESSIONS

- Say, “Look at all of these pictures.”
- Say, “There are eight pictures in all.”
- Say, “Listen carefully while I tell you what each picture is.”
- In random order, read the titles of the pictures to the child and place each picture in front of the child on the table in a horizontal row.
- Say, “I am going to use these pictures to make a thinking map about sandwiches.”
- Say, “Watch and listen to what I do and say.”
- Take Pictures 1 and 7 and place them on one side of the white board approximately 4” apart.
- Draw a line between the pictures with the white board marker.
- Say, “For lunch, I often eat peanut butter and jelly sandwiches.”
- Take Picture 6 and place it on the white board approximately 4” away from Picture 7.
- Draw a line between Pictures 6 and 7 with the white board marker.
- Say, “When I was a teacher, I noticed that many of my students also ate peanut butter and jelly sandwiches for lunch.”
- Take Picture 4 and place it on the white board approximately 4” away from Picture 1.
- Draw a line between Pictures 1 and 4 with the white board marker.
- Say, “When I make peanut butter and jelly sandwiches, I use a knife like this.”
- Take Picture 2 and place it on the white board approximately 4” away from Picture 4.
- Draw a line between Pictures 2 and 4 with the white board marker.
- Say, “I like a lot of peanut butter on my sandwiches so I use the knife to spread a big blob of peanut butter on the bread.”
- Take Picture 3 and place it on the white board approximately 4” away from Picture 4.
- Draw a line between Pictures 3 and 4 with the white board marker.
- Say, “I like just a little bit of jelly on my sandwiches so I use the knife to spread a small blob of jelly on the bread.”
- Say, “Ok, my thinking map about sandwiches is finished. I don’t have any ideas for using these two pictures in my map.”
- Point to Pictures 5 and 8.
- Say, “Do you have any ideas for using these two pictures?”
- Child’s anticipated response: “You make [sandwiches, lunch, food] in the kitchen.”
- Say, “Yes, that’s right. Great idea! I make my peanut butter and jelly sandwiches for lunch in the kitchen. I think that I need to make some changes and put this picture in my thinking map. Watch and listen to what I do and say.”
- Take Pictures 2, 3, and 4 off the white board and erase the lines.

- Say, “I am going to take these pictures down for a minute and get rid of these lines.”
- Take Picture 5 and place it approximately 4” from Picture 1.
- Draw a line between Pictures 1 and 5 with the white board marker.
- Say, “I make peanut butter and jelly sandwiches for lunch in the kitchen.”
- Take Picture 4 and place it approximately 4” from Picture 5.
- Draw a line between Pictures 4 and 5 with the white board marker.
- Say, “I get the knife to make my sandwich out of the drawer in the kitchen.”
- Take Pictures 2 and 3 and place them approximately 4” from Picture 4.
- Draw a line between Pictures 2 and 4with the white board marker.
- Draw a line between Pictures 3 and 4 with the white board marker.
- Say, “I use the knife to spread a big blob of peanut butter and a little blob of jelly on the bread just like I like.”

Finished thinking map:



* Picture 8 not used intentionally

APPENDIX H
MAP SCORING SHEET

Child #: Pre / Post (circle one) Rater #:

MAPPING SESSION DATA SECTION			RATER SCORING SECTION				
			Scoring Systems 1 and 2		**Scoring System 2**		
Column 1 Proposition # (in order created by the child)	*Column 2* Concept links (pairs of pictures)	*Column 3* Child's verbal explanations for concept links	*Column 4* Concept links representing accurate relationships from proposition inventory (Y/N)	*Column 5* Accurate verbal explanations for concept links (Y/N/NA)	*Column 6* Complete verbal explanations for concept links (Y/N/NA)	*Column 7* Concept links representing critical links from proposition inventory (Y/N)	*Column 8* Accurate verbal explanations for critical links (Y/N/NA)
			Total=	Total=	Total=		Total=

Scoring system 1

A. ______(Column 4 total) * 1 pt = _______pts
B. ______(Column 5 total) * 1 pt = _______pts
C. Total = _______pts

Scoring system 2

A. ______(Column 4 total) * 1 pt = _______pts
B. ______(Column 5 total) * 1 pt = _______pts
C. ______(Column 6 total) * 1 pt = _______pts
D. ______(Column 8 total) * 1 pt = _______ pts
E. Total = _______pts

Scoring system 3

SCORE POINT VALUE (CIRCLE ONE)	SCORING CRITERIA
4-Advanced	At least 80% of the propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. At least two critical links are represented accurately.
3-Proficient	At least 60% of the propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. At least one critical link is represented accurately.
2-Basic	At least 40% of the propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. At least one critical link is represented.
1-Beginning	Less than 40% propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. None of the critical links are represented.

Insert picture of map here

APPENDIX I
USE OF MAP SCORING SHEET

Preparation of Map Scoring Sheets:

1. Record the child's participant number at the top of the map scoring sheet.
2. Circle the mapping session, pre or post, at the top of the map scoring sheet.
3. Watch the videotape of the mapping session and enter the following information about each proposition in the three columns of the section labeled "Mapping Session Data Section."
   a. Column 1: Record number of proposition (in order created by the child)
   b. Column 2: List names of pictures paired to form concept links (e.g., Butterfly-Metamorphosis)
   c. Column 3: Transcribe conversation segment between the child and interviewer about each concept link
4. Insert a picture of the completed map at the bottom of the map scoring sheet.

Directions for Raters:

1. Enter your individual rater number at the top of the map scoring sheet.
2. Evaluate each proposition on the map scoring sheet individually for both scoring systems one and two.
3. For **scoring system one**, enter the following information about each proposition in the first two columns of the section labeled "Rater Scoring Section."
   a. Column 4: Refer to the proposition inventory. Do the concept links (pairs of pictures) match any accurate relationships in the proposition inventory? Enter Y for "Yes" or N for "No."
   b. Column 5: For concept links that represented accurate relationships (Yes in Column 4), read the transcribed conversation segment between the interviewer and the child. Is the child's explanation about the relationship between the pair of pictures scientifically accurate? Enter Y for "Yes" or N for "No." For inaccurate relationships (No in Column 5), enter NA for "Not applicable."
   c. Total Columns 4 and 5 and enter these two totals in the first worksheet below the table.
   d. Find the sum of parts A and B in the worksheet to find the total points assigned using scoring system one.
4. For **scoring system two**, enter the following information about each proposition in the last three columns of the section labeled "Rater Scoring Section."
   a. Column 6: For concept links that represent accurate relationships and are explained accurately (Yes in both Columns 4 and 5), re-read the transcribed conservation segment between the interviewer and the child. Is the child's explanation about the relationship between the pair of

pictures complete? Enter Y for “Yes” or N for “No.” For inaccurate relationships between concept links or accurate relationships between concept links lacking accurate explanations, enter NA for “Not applicable.”

   b. Column 7: For concept links that represent accurate relationships (Yes in Column 4), refer again to the proposition inventory. Do the concept links match any of the five critical links? Enter Y for “Yes” or N for “No.”
   c. Column 8: Is the child’s explanation about the relationship between the pair of pictures a scientifically accurate critical link (Yes in Columns 4, 5, and 7)? Enter Y for “Yes” or N for “No.” For non-critical links, enter NA for “Not applicable.”
   d. Enter the totals for Columns 4 and 5 in the second worksheet below the table.
   e. Total Columns 6 and 8 and enter these totals in the worksheet.
   f. Find the sum of parts A, B, C, and D in the worksheet to find the total points assigned using scoring system two

5. For **scoring system three**, review all of the propositions on the map scoring sheet.
6. Use the scoring criteria in the rubric to determine the level of performance and points assigned for the map. The four levels of performance differ in how well the propositions in the map represent a) accurate and complete understanding of butterflies and their role in the ecosystem, and b) critical links.

APPENDIX J
EXAMPLE MAP SCORING SHEET BEFORE SCORING

Child #: 2 [✓ Pre]/Post Rater #:

MAPPING SESSION DATA SECTION			RATER SCORING SECTION				
			Scoring Systems 1 and 2		Scoring System 2		
Column 1 Proposition # (in order created by the child)	*Column 2* Concept links (pairs of pictures)	*Column 3* Child's verbal explanations for concept links	*Column 4* Concept links representing accurate relationships from proposition inventory (Y/N)	*Column 5* Accurate verbal explanations for concept links (Y/N/NA)	*Column 6* Complete verbal explanations for concept links (Y/N/NA)	*Column 7* Concept links representing critical links from proposition inventory (Y/N)	*Column 8* Accurate verbal explanations for critical links (Y/N/NA)
1	Butterfly/Biodiversity	C-I think that this one should go here and this one should go here because this is one butterfly and these are many I-So that's one type of butterfly you're saying C-That's one type and those are a bunch of types					
2	Butterfly/Rainforest layers	C-And they [butterflies] can live in the rainforest					
3,4	Animal-animal interactions/Rainforest layers	C-And there can be predators like a praying mantis in there [the rainforest]					
	Animal-animal interactions/Butterfly						
5	Butterfly/Metamorphosis	C- It [the butterfly] can grow up and migrate.					
6,7	Biodiversity/Metamorphosis	C-It's a butterfly but it can turn into any kind in the rainforest					
	Rainforest layers/Metamorphosis						
8	Rainforest layers/Conservation	C-It [the butterfly] can live in the wild or here. *Child points to the Conservation and Rainforest layers pictures.*					
9	Conservation/Metamorphosis	C-It [the butterfly] can grow up in this one or this one *Child points to the Conservation and Rainforest*					

		layers pictures.					
			Total=	Total=	Total=		Total=

Scoring system 1

D. ______(Column 4 total) * 1 pt = _______pts
E. ______(Column 5 total) * 1 pt = _______pts
F. Total = _______pts

Scoring system 2

F. ______(Column 4 total) * 1 pt = _______pts
G. ______(Column 5 total) * 1 pt = _______pts
H. ______(Column 6 total) * 1 pt = _______pts
I. ______(Column 8 total) * 1 pt = _______ pts
J. Total = _______pts

Scoring system 3

SCORE POINT VALUE (CIRCLE ONE)	SCORING CRITERIA
4-Advanced	At least 80% of the propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. At least two critical links are represented accurately.
3-Proficient	At least 60% of the propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. At least one critical link is represented accurately.
2-Basic	At least 40% of the propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. At least one critical link is represented.
1-Beginning	Less than 40% propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. None of the critical links are represented.

APPENDIX K
EXAMPLE MAP SCORING SHEET AFTER SCORING

Child #: 2 [✓ Pre]/Post Rater #: 1

MAPPING SESSION DATA SECTION			RATER SCORING SECTION				
			Scoring Systems 1 and 2		Scoring System 2		
Column 1 Proposition # (in order created by the child)	*Column 2* Concept links (pairs of pictures)	*Column 3* Child's verbal explanations for concept links	*Column 4* Concept links representing accurate relationships from proposition inventory (Y/N)	*Column 5* Accurate verbal explanations for concept links (Y/N/NA)	*Column 6* Complete verbal explanations for concept links (Y/N/NA)	*Column 7* Concept links representing critical links from proposition inventory (Y/N/NA)	*Column 8* Accurate verbal explanations for critical links (Y/N/NA)
1	Butterfly/Biodiversity	C-I think that this one should go here and this one should go here because this is one butterfly and these are many I-So that's one type of butterfly you're saying C-That's one type and those are a bunch of types	Y	Y	Y	N	NA
2	Butterfly/Rainforest layers	C-And they [butterflies] can live in the rainforest	Y	N	N	N	NA
3,4	Animal-animal interactions/Rainforest layers	C-And there can be predators like a praying mantis in there [the rainforest]	Y	N	N	N	NA
	Animal-animal interactions/Butterfly		Y	Y	Y	N	NA
5	Butterfly/Metamorphosis	C- It [the butterfly] can grow up and migrate.	Y	N	N	N	NA
6,7	Biodiversity/Metamorphosis	C-It's a butterfly but it can turn into any kind in the rainforest	N	NA	NA	NA	NA
	Rainforest layers/Metamorphosis		N	NA	NA	NA	NA
8	Rainforest layers/Conservation	C-It [the butterfly] can live in the wild or here. *Child points to the Conservation and Rainforest layers pictures.*	Y	N	N	N	NA
9	Conservation/Metamorphosis	C-It [the butterfly] can grow up in this one or this one *Child points to the Conservation and Rainforest*	N	NA	NA	N	NA

		layers pictures.					
			Total=6	Total= 2	Total=2		Total=0

Scoring system 1

G. 6 (Column 4 total) * 1 pt = 6 pts
H. 2 (Column 5 total) * 1 pt = 2 pts
I. Total = 8 pts

Scoring system 2

K. 6 (Column 4 total) * 1 pt = 6 pts
L. 2 (Column 5 total) * 1 pt = 2 pts
M. 2 (Column 6 total) * 1 pt = 2 pts
N. 0 (Column 8 total) * 1 pt = 0 pts
O. Total = 10 pts

Scoring system 3

SCORE POINT VALUE (CIRCLE ONE)	SCORING CRITERIA
4-Advanced	At least 80% of the propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. At least two critical links are represented accurately.
3-Proficient	At least 60% of the propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. At least one critical link is represented accurately.
2-Basic	At least 40% of the propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. At least one critical link is represented.
[✓ 1-Beginning]	Less than 40% propositions on the map reflect accurate and complete scientific understanding of butterflies and their role in ecosystems. None of the critical links are represented.

APPENDIX L
TALLY OF ALL SCORES GENERATED USING SCORING SYSTEM ONE

Occasion	Child	Rater 1	Rater 2	Rater 3
Pre-visit	1	6	6	5
	2	8	8	8
	3	7	7	7
	4	8	8	8
	5	7	7	9
	6	6	6	6
	7	3	3	3
	8	16	16	16
	9	7	7	9
	10	9	7	8
	11	5	5	3
	12	15	15	15
	13	5	4	6
	14	4	2	5
	15	3	3	3
	16	6	5	6
	17	9	8	9
	18	5	5	6
	19	5	5	4
	20	4	4	4
	21	12	12	13
	22	13	13	10
	23	13	13	12
	24	9	9	9
	25	9	10	9
	26	10	8	11
	27	7	7	7
	28	5	5	5
	29	3	3	3
	30	18	18	18
	31	5	5	4
	32	3	3	3
	33	7	7	7
	34	5	5	4
	35	2	2	2
	36	6	7	7
	37	9	9	10
	38	10	11	10
	39	6	6	6
	40	7	8	6
	41	10	8	9
	42	9	9	9

Occasion	Child	Rater 1	Rater 2	Rater 3
Post	1	7	7	7
	2	10	10	10
	3	12	12	10
	4	9	9	8
	5	7	7	9
	6	6	6	6
	7	4	4	5
	8	16	16	18
	9	11	11	10
	10	6	7	6
	11	6	5	6
	12	14	15	14
	13	7	5	7
	14	6	7	7
	15	6	5	7
	16	5	5	8
	17	10	10	12
	18	6	6	7
	19	8	8	7
	20	5	5	5
	21	12	12	13
	22	0	0	0
	23	18	18	18
	24	5	4	5
	25	9	10	8
	26	9	7	9
	27	7	7	7
	28	5	5	4
	29	3	3	3
	30	19	19	19
	31	6	6	5
	32	8	9	7
	33	10	8	10
	34	5	5	4
	35	7	7	6
	36	10	11	11
	37	10	10	11
	38	5	6	4
	39	8	9	8
	40	7	8	6
	41	5	5	6
	42	9	9	9

APPENDIX M
TALLY OF ALL SCORES GENERATED USING SCORING SYSTEM TWO

Occasion	**Child**	**Rater 1**	**Rater 2**	**Rater 3**
Pre-visit	1	7	7	5
	2	11	11	11
	3	7	8	7
	4	8	8	8
	5	7	7	9
	6	8	8	7
	7	5	4	4
	8	20	20	20
	9	10	9	12
	10	8	8	10
	11	6	6	4
	12	17	18	17
	13	8	6	9
	14	4	2	6
	15	4	4	4
	16	6	6	6
	17	8	9	10
	18	5	5	6
	19	5	5	4
	20	4	5	4
	21	13	13	13
	22	13	15	10
	23	13	18	12
	24	10	11	10
	25	10	13	11
	26	10	9	11
	27	7	7	7
	28	7	7	5
	29	3	3	3
	30	20	21	18
	31	5	5	4
	32	3	3	3
	33	9	10	9
	34	7	7	4
	35	2	2	2
	36	8	8	7
	37	9	9	10
	38	11	12	10
	39	6	6	6
	40	8	9	6
	41	10	8	10
	42	10	11	10

Occasion	Child	Rater 1	Rater 2	Rater 3
Post-visit	1	7	8	7
	2	13	13	13
	3	11	15	11
	4	11	12	10
	5	7	7	9
	6	8	8	7
	7	5	9	7
	8	20	19	22
	9	15	15	13
	10	8	9	6
	11	6	6	7
	12	17	19	16
	13	7	7	10
	14	8	8	7
	15	7	6	10
	16	7	6	10
	17	10	12	14
	18	6	6	7
	19	7	8	7
	20	6	7	6
	21	12	13	13
	22	0	0	0
	23	18	25	18
	24	5	4	5
	25	10	13	10
	26	10	10	9
	27	7	7	7
	28	6	7	4
	29	3	3	3
	30	20	23	19
	31	6	6	5
	32	9	10	7
	33	12	12	13
	34	7	7	4
	35	8	8	6
	36	11	10	11
	37	10	12	11
	38	6	8	4
	39	8	9	8
	40	8	9	6
	41	6	5	7
	42	10	11	10

APPENDIX N
TALLY OF ALL SCORES GENERATED USING SCORING SYSTEM THREE

Occasion	Child	Rater 1	Rater 2	Rater 3
Pre-visit	1	1	1	1
	2	1	1	1
	3	1	1	1
	4	1	1	1
	5	1	1	1
	6	1	1	1
	7	1	1	1
	8	2	1	3
	9	2	1	2
	10	2	1	2
	11	1	1	1
	12	2	1	2
	13	1	1	1
	14	1	1	1
	15	1	1	1
	16	1	1	1
	17	2	1	2
	18	1	1	1
	19	1	1	1
	20	1	1	1
	21	2	1	2
	22	2	1	1
	23	2	1	2
	24	2	1	2
	25	1	1	1
	26	1	1	1
	27	1	1	1
	28	1	1	1
	29	1	1	1
	30	1	1	1
	31	1	1	1
	32	1	1	1
	33	2	1	2
	34	1	1	1
	35	1	1	1
	36	1	1	1
	37	1	1	1
	38	1	1	1
	39	1	1	1
	40	1	1	1
	41	1	1	1
	42	1	1	1

Occasion	**Child**	**Rater 1**	**Rater 2**	**Rater 3**
Post	1	1	1	1
	2	1	1	1
	3	1	1	1
	4	1	1	1
	5	1	1	1
	6	1	1	1
	7	2	1	2
	8	2	1	3
	9	2	1	2
	10	1	1	1
	11	1	1	1
	12	1	1	2
	13	1	1	1
	14	1	1	1
	15	2	1	3
	16	1	1	1
	17	2	1	2
	18	1	1	1
	19	1	1	1
	20	1	1	1
	21	2	1	2
	22	1	1	1
	23	2	1	2
	24	1	1	1
	25	1	1	1
	26	1	1	1
	27	1	1	1
	28	1	1	1
	29	1	1	1
	30	1	1	1
	31	1	1	1
	32	1	1	1
	33	2	1	3
	34	1	1	1
	35	1	1	1
	36	1	1	1
	37	1	1	1
	38	1	1	1
	39	1	1	1
	40	1	1	1
	41	1	1	1
	42	1	1	1

APPENDIX O
INITIAL CODES MAPPED TO THE SIX CATEGORIES OF VERBAL EXPLANATIONS

Initial codes	**Categories**
Feeding on plants Feeding on animals Need for shelter Need for water Need for air Need for sunlight	Needs of butterflies
Changes from egg to larva Changes from larva to pupa Changes from pupa to adult Changes in size	Life cycle of butterflies
Predation of butterflies Predation of other insects Butterfly defenses Butterfly habitats Other insect habitats Insect helpers Migration Niches	Ecology of butterflies and other insects
Butterfly body parts Insect body parts Characteristics of insects Characteristics of spiders Locomotion Diversity of butterflies Diversity of insects Personal experiences	Diversity and classification of butterflies and other insects
People harming butterflies People helping butterflies Habitat destruction “Artificial” habitats Endangered species Extinction	Threats to and conservation of butterflies
Emotions of butterflies Butterfly mothers Butterfly friends Comparing butterflies to bees	Social behavior of butterflies

LIST OF REFERENCES

Allen, S. (1997). Using scientific inquiry activities in exhibit explanations. *Science Education, 81*(6), 715-734.

Allen, S. (2002). Looking for learning in visitor talk: A methodological exploration. In G. Leinhardt, K. Crowley & K. Knutson (Eds.), *Learning conversations in museums* (pp. 259-304). Mahwah, NJ: Lawrence Erlbaum Associates.

Allen, S. (2004). Designs for learning: Studying science museum exhibits that do more than entertain. *Science Education, 88*(S1), S17-S33.

Allen, S. (2008). Evaluating exhibitions In A. Friedman (Ed.), *Framework for evaluating impacts of informal science education projects* (pp. 45-59). Washington, DC: National Science Foundation

Anderson, D., & Lucas, K. B. (1997). The effectiveness of orienting students to the physical features of a science museum prior to visitation. *Research in Science Education, 27*(4), 485-495.

Anderson, D., Lucas, K. B., Ginns, I. S., & Dierking, L. D. (2000). Development of knowledge about electricity and magnetism during a visit to a science museum and related post-visit activities. *Science Education, 84*(5), 658-679.

Ash, D. (2003). Dialogic inquiry in life science conversations of family groups in a museum. *Journal of Research in Science Teaching, 40*(2), 138-162.

Ash, D. (2004). Reflective scientific sense-making dialogue in two languages: The science in the dialogue and the dialogue in the science. *Science Education, 88*(6), 855-884.

Ash, D. (2007). Talk, tools, and tensions: Observing biological talk over time. *International Journal of Science Education, 29*(12), 1581-1602.

American Association for the Advancement of Science. (1993). Benchmarks for scientific literacy. New York, NY: Oxford University Press.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). The standards for educational and psychological testing. Washington, DC: AERA Publications.

Association of Children's Museums. (2009). *Stats and trends*. Retrieved March 19, 2009, from http://www.childrensmuseums.org/about/facts.htm.

Association of Science and Technology Centers. (2009). *Science center highlights*. Retrieved March 19, 2009, from http://www.astc.org/about/pdf/Backgrounders/Highlights2009.pdf.

Ausubel, D. P. (1960). The use of advance organizers in learning and retention of meaningful information. *Journal of Educational Psychology, 51*(5), 267-272.

Barrow, L. H. (2002). What do elementary students know about insects? *Journal of Elementary Science Education, 14*(2), 51-56.

Bell, P., Lewenstein, B., Shouse, A.W., & Feder, M.A. (Eds.). (2009). *Learning science in informal environments: Places, people, and pursuits.* Washington, DC: The National Academies Press.

Blud, L. M. (1990a). Social interaction and learning among family groups visiting a museum. *Museum Management and Curatorship, 9*(1), 43-51.

Blud, L. M. (1990b). Sons and daughters; observations on the way families interact during a museum visit. *Museum Management and Curatorship, 9*(3), 257-264.

Bogdan, R., & Biklen, S. (1998). *Qualitative Research for Education*. Boston, MA: Allyn and Bacon.

Boisvert, D. L., & Slez, B. J. (1994). The relationship between visitor characteristics and learning-associated behaviors in a science museum discovery space. *Science Education, 78*(2), 503-518.

Borun, M., Chambers, M., & Cleghorn, A. (1996). Families are learning in science museums. *Curator, 39*(2), 123-138.

Borun, M., Chambers, M., Dritsas, J., & Johnson, J. (1997). Enhancing family learning through exhibits. *Curator, 40*(4), 279-295.

Borun, M., & Dritsas, J. (1997). Developing family-friendly exhibits. *Curator, 40*(3), 178-196.

Bredekamp, S., & Rosegrant, T. (1995). *Reaching potentials: Transforming early childhood curriculum and assessment*. Washington, DC: National Association for the Education of Young Children.

Britsch, S. (2001). Assessment of emergent science literacy in classrooms for young children. In D. Shepardson (Ed.), *Assessment in science* (pp. 67-82). Dordrecht, The Netherlands: Kluwar Academic Publishers.

Brown, C. (1995). Making the most of family visits: Some observations of parents with children in a museum science centre. *Museum Management and Curatorship, 14*(1), 65-71.

Burkham, D., Ready, D., Lee, V., & LoGerfo, L. (2003). *Social class differences in summer learning between kindergarten and first grade*. Ann Arbor, MI: University of Michigan.

Carrier, S. J. (2009). Environmental education in the schoolyard: Learning styles and gender. *The Journal of Environmental Education*, *40*(3), 2-12.

Chermayeff, J. C., Blandford, R. J., & Losos, C. M. (2001). Working at play: Informal science education on museum playgrounds. *Curator, 44*(1), 47-60.

Chrispeels, J. (1996). Effective schools and home-school-community partnership roles: A framework for parent involvement. *School Effectiveness and School Improvement, 7*(4), 297-323.

Cosmos Corporation. (1998). A report on the evaluation of the National Science Foundation's Informal Science Education program [Electronic Version]. Retrieved March 19, 2009, from http://www.nsf.gov/pubs/1998/nsf9865/nsf9865.htm.

Cox-Petersen, A. M., Marsh, D. D., Kisiel, J., & Melber, L. M. (2003). Investigation of guided school tours, student learning, and science reform recommendations at a museum of natural history. *Journal of Research in Science Teaching, 40*(2), 200-218.

Cox-Petersen, A. M., & Pfaffinger, J. A. (1998). Teacher preparation and teacher-student interactions at a discovery center of natural history. *Journal of Elementary Science Education, 10*(2), 20-35.

Cronin-Jones, L. (2000). The effectiveness of schoolyards as sites for elementary science instruction. *School Science and Mathematics,* 100(4), 203-11.

Crowley, K., Callanan, M. A., Jipson, J. L., Galco, J., Topping, K., & Shrager, J. (2001). Shared scientific thinking in everyday parent-child activity. *Science Education, 85*(6), 712-732.

Crowley, K., Callanan, M. A., Tenenbaum, H. R., & Allen, E. (2001). Parents explain more often to boys than to girls during shared scientific thinking. *Psychological Science, 12*(3), 258-261.

Crowley, K., & Jacobs, M. (2002). Building islands of expertise in everyday family activity. In G. Leinhardt, K. Crowley & K. Knutson (Eds.), *Learning conversations in museums* (pp. 305-330). Mahwah, NJ: Lawrence Erlbaum Associates.

Davis, J. (2009). Revealing the research ‘hole’ of early childhood education for sustainability: A preliminary survey of the literature. *Environmental Education Research, 15*(2), 227-241.

Dierking, L. D., Falk, J. H., Rennie, L., Anderson, D., & Ellenbogen, K. (2003). Policy statement of the" Informal Science Education" ad hoc committee. *Journal of Research in Science Teaching, 40*(2), 108-111.

Dillon, J., Rickinson, M., Teamey, K., Morris, M., Choi, M. Y., Sanders, D., et al. (2006). The value of outdoor learning: evidence from research in the UK and elsewhere. *School Science Review, 87*(320), 107.

Driver, R., Asoko, H., Leach, J., Scott, P., & Mortimer, E. (1994). Constructing scientific knowledge in the classroom. *Educational researcher, 23*(7), 5-12.

Duschl, R.A., Schweingruber, H.A., & Shouse, A.W. (Eds.). (2007). *Taking science to school: Learning and teaching science in Grades K-8*. Washington, DC: National Academies Press.

Ellenbogen, K. (2002). Museums in family life: An ethnographic case study. In G. Leinhardt, K. Crowley & K. Knutson (Eds.), *Learning conversations in museums* (pp. 81-102). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Ellenbogen, K., Luke, J., & Dierking, L. D. (2004). Family learning research in museums: An emerging disciplinary matrix? *Science Education, 88*(S1), S48-S58.

Falk, J. H. (1982). The field trip milieu: Learning and behavior as a function of contextual events. *Journal of Educational Research, 76*(1), 22-28.

Falk, J. H. (1997). Testing a museum exhibition design assumption: Effect of explicit labeling of exhibit clusters on visitor concept development. *Science Education, 81*(6), 679-687.

Falk, J. H., & Adelman, L. M. (2003). Investigating the impact of prior knowledge and interest on aquarium visitor learning. *Journal of Research in Science Teaching, 40*(2), 163-176.

Falk, J. H., & Dierking, L. D. (1992). *The museum experience*. Washington, DC: Whalesback Books.

Falk, J. H. & Dierking, L. D. (2000). *Learning from museums: Visitor experiences and the making of meaning.* Walnut Creek, CA: Alta Mira Press.

Falk, J. H., Moussouri, T., & Coulson, D. (1998). The effect of visitors' agendas on museum learning. *Curator, 41*(2), 107-120.

Falk, J. H., & Storksdieck, M. (2005). Using the contextual model of learning to understand visitor learning from a science center exhibition. *Science Education, 89*(5), 744-778.

Fender, J. G., & Crowley, K. (2007). How parent explanation changes what children learn from everyday scientific thinking. *Journal of Applied Developmental Psychology, 28*(3), 189-210.

Figueiredo, M., Lopes, A. S., Firmino, R., & de Sousa, S. (2004, April). *"Things we know about the cow": Concept mapping in a preschool setting.* Paper presented at the the First International Conference on Concept Mapping, Pamploma, Spain.

Fleer, M. (1996). Early learning about light: Mapping preschool children's thinking about light before, during and after involvement in a two week teaching program. *International Journal of Science Education, 18*(7), 819-836.

Florida Department of Education. (1996). *Sunshine state standards*. Retrieved March 19, 2009, from http://www.fldoe.org/bii/curriculum/sss/sss1996.asp.

Gallegos, L. (1994). Preconceptions and relations used by children in the construction of food chains. *Journal of Research in Science Teaching, 31*(3), 259-272.

Gallenstein, N. (2003). *Creative construction of mathematics and science concepts in early childhood*. Olney, MD: Association for Childhood Education International.

Gilbert, J., Osborne, R. J., & Fensham, P. J. (1982). Children's science and its consequences for teaching. *Science Education, 66*(4), 623-633

Gilbert, J., & Priest, M. (1997). Models and discourse: A primary school science class visit to a museum. *Science Education, 81*, 749-762.

Gleason, M. E., & Schauble, L. (1999). Parents' assistance of their children's scientific reasoning. *Cognition and Instruction, 17*(4), 343-378.

Greenfield, T. A. (1995). Sex differences in science museum exhibit attraction. *Journal of Research in Science Teaching, 32*(9), 925-938.

Hatano, G., & Inagaki, K. (1995). Young children's naive theory of biology. In J. Mehler & S. Franck (Eds.), *Cognition on Cognition* (pp. 153-170). Boston, MA: MIT Press.

Hilke, D. D. (1989). The family as a learning system: An observational study of families in museums. In B. Butler & S. Sussman (Eds.), *Museum visits and activities for family life enrichment* (pp. 3-33). New York, NY: Haworth Press.

Hood, M. G. (2004). Staying away: Why people choose not to visit museums. In G. Anderson (Ed.), *Reinventing the Museum.* Walnut Creek, CA: Rowman Altamira.

Hunter, J., Monroe-Ossi, H., & Fountain, C. (2008, April). *Young Florida naturalists: Concept mapping and learning of preschool children.* Paper presented at the Third International Conference on Concept Mapping, Talinn, Estonia.

James, J. J., & Bixler, R. D. (2008). Children's role in meaning making through their participation in an environmental education program. *The Journal of Environmental Education, 39*(4), 44-59.

Johnson, J., & Onwuegbuzie, A. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14-26.

Johnson, K. E., Alexander, J. M., Spencer, S., Leibham, M. E., & Neitzel, C. (2004). Factors associated with the early emergence of intense interests within conceptual domains. *Cognitive Development, 19*(3), 325-343.

Kelemen, D., Callanan, M. A., Casler, K., & Pérez-Granados, D. R. (2005). Why things happen: Teleological explanation in parent–child conversations. *Developmental Psychology, 41*(1), 251-264.

Klein, D. C. D., Chung, G., Osmundson, E., Herl, H. E., & O'Neil Jr, H. F. (2002). *Examining the validity of knowledge mapping as a measure of elementary students' scientific understanding*. Retrieved March 19, 2009, from University of California, Los Angeles Center for the Study of Evaluation Web site: http://research.cse.ucla.edu/Reports/TR557.pdf.

Korpan, C. A., Bisanz, G. L., Bisanz, J., Boehme, C., & Lynch, M. A. (1997). What did you learn outside of school today? Using structured interviews to document home and community activities related to science and technology. *Science Education, 81*(6), 651-662.

Kubota, C. A., & Olstad, R. G. (1991). Effects of novelty-reduced preparation on exploratory behavior and cognitive learning in a science museum setting. *Journal of Research in Science Teaching, 28*(3), 225-234.

Leach, J., Driver, R., Scott, P., & Wood-Robinson, C. (1995). Children's ideas about ecology 1: theoretical background, design and methodology. *International Journal of Science Education, 17*(6), 721-732.

Leach, J., Driver, R., Scott, P., & Wood-Robinson, C. (1996a). Children's ideas about ecology 2: ideas found in children aged 5-16 about the cycling of matter. *International Journal of Science Education, 18*(1), 19-34.

Leach, J., Driver, R., Scott, P., & Wood-Robinson, C. (1996b). Children's ideas about ecology 3: ideas found in children aged 5-16 about the interdependency of organisms. *International Journal of Science Education, 18*(2), 129-141.

Lind, K. K. (2005). *Exploring science in early childhood education*. Albany, NY: Delmar Thompson Learning.

Littledyke, M. (2008). Science education for environmental awareness: Approaches to integrating cognitive and affective domains. *Environmental Education Research, 14*(1), 1-17.

McClure, J.R., & Bell, P.E. (1990). Effects of an environmental education related STS approach instruction on cognitive structures of pre-service science teachers. University Park, PA: Pennsylvania State University. (ERIC Document Reproduction Services No. ED 341 582).

McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching, 36*(4), 475-492.

McManus, P. (1987). It's the company you keep: The social determination of learning-related behavior in a science museum. *International Journal of Museum Management and Curatorship, 6*, 263-270.

Melton, A. W. (1936). *Experimental studies of the education of children in a museum of science.* Washington, DC: Publications of the American Association of Museums.

Miller, J. D. (2004). Public understanding of, and attitudes toward, scientific research: What we know and what we need to know. *Public Understanding of Science, 13*(3), 273.

Mindes, G. (2003). *Assessing young children*. Upper Saddle River, NJ: Merrill Prentice Hall.

Monroe-Ossi, H., Wehry, S., Algina, J., & Hunter, J. (2008, April). *Healthy habits through literacy: A concept mapping and health curriculum for preschool and prekindergarten children.* Paper presented at the Third International Conference on Concept Mapping, Talin, Estonia.

Myers, O. E., Saunders, C. D., & Garrett, E. (2004). What do children think animals need? Developmental trends. *Environmental Education Research, 10*(4), 545-562.

National Science Foundation. (2007). *Award abstract #0638990*. Retrieved March 19, 2009, from http://nsf.gov/awardsearch/showAward.do?AwardNumber=0638990.

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

Novak, J. D. (2005). Results and implications of a 12-year longitudinal study of science concept learning. *Research in Science Education, 35*(1), 23-40.

Novak, J.D., & Canas, A.J. (2007). *The theory underlying concept maps and how to construct and use them.* Retrieved March 19, 2009, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.137.2955&rep=rep1&type=pdf.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge, MA: Cambridge University Press.

O'Byrne, B. (2008). Knowing more than words can say: Using multimodal assessment tools to excavate and construct knowledge about wolves. *International Journal of Science Education, 31*(4), 1-17.

Packer, J., & Ballantyne, R. (2005). Solitary vs. shared learning: Exploring the social dimension of museum learning. *Curator, 48*(2), 177-192.

Palmquist, S., & Crowley, K. (2007). From teachers to testers: How parents talk to novice and expert children in a natural history museum. *Science Education, 91*(5), 783-804.

Piscitelli, B., & Anderson, D. (2001). Young children's perspectives of museum settings and experiences. *Museum Management and Curatorship, 19*(3), 269-282.

Prokop, P., Prokop, M., & Tunnicliffe, S. D. (2007). Effects of keeping animals as pets on children's concepts of vertebrates and invertebrates. *International Journal of Science Education, 30*(4), 431-449.

Rennie, L. (2007). Learning science outside of school. In S. Abell & N. Lederman (Eds.), *Handbook of research on science education* (pp. 125-157). Mahwah, NJ: Lawrence Erlbaum Associates.

Rennie, L., Feher, E., Dierking, L. D., & Falk, J. H. (2003). Toward an agenda for advancing research on science learning in out-of-school settings. *Journal of Research in Science Teaching, 40*(2), 112-120.

Rennie, L., & Johnston, D. J. (2004). The nature of learning and its implications for research on learning from museums. *Science Education, 88*(S1), S4-S16.

Rice, D. C., Ryan, J. M., & Samson, S. M. (1998). Using concept maps to assess student learning in the science classroom: Must different methods compete? *Journal of Research in Science Teaching, 35*(10), 1103-1127.

Ruiz-Primo, M. A., Schultz, S. E., & Shavelson, R. J. (1997). *Concept map-based assessment in science: Two exploratory studies.* Retrieved March 19, 2009, from University of California, Los Angeles Center for Research on Evaluation, Standards, and Student Testing Web site: http://www.cse.ucla.edu/products/Reports/TECH436.pdf.

Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching, 33*(6), 569-600.

Ruiz-Primo, M. A., Shavelson, R. J., Li, M., & Schultz, S. E. (2001). On the validity of cognitive interpretations of scores from alternative concept-mapping techniques. *Educational Assessment, 7*(2), 99-141.

Rye, J. A., & Rubba, P. A. (1998). An exploration of the concept map as an interview tool to facilitate the externalization of students' understandings about global atmospheric change. *Journal of Research in Science Teaching, 35*(5), 521-546.

Sandifer, C. (1997). Time-based behaviors at an interactive science museum: Exploring the differences between weekday/weekend and family/nonfamily visitors. *Science Education, 81*(6), 689-701.

Schau, C., Mattern, N., Zeilik, M., Teague, K. W., & Weber, R. J. (2001). Select-and-fill-in concept map scores as a measure of students' connected understanding of science. *Educational and Psychological Measurement, 61*(1), 136-146.

Schauble, L., & Bartlett, K. (1997). Constructing a science gallery for children and families: The role of research in an innovative design process. *Science Education, 81*(6), 781-793.

Shavelson, R. J., & Ruiz-Primo, M. A. (2000). On the psychometrics of assessing science understanding. In J. Mintzes, J. H. Wandersee, & J. D. Novak (Eds.), *Assessing science understanding* (pp. 304-341). San Diego, CA: Academic Press.

Shepardson, D.P. (1997). Of butterflies and beetles: First graders' ways of seeing and talking about insect life cycles. *Journal of Research in Science Teaching, 34*(9), 873-889.

Shepardson, D. P. (2002). Bugs, butterflies, and spiders: Children's understandings about insects. *International Journal of Science Education, 24*(6), 627-643.

Shine, S., & Acosta, T. (2000). Parent-child social play in a children's museum. *Family Relations, 49*(1), 45-52.

Shonkoff, J. P., & Phillips, D. (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.

Siegal, M. (1999). Language and thought: The fundamental significance of conversational awareness for cognitive development. *Developmental Science, 2*(1), 1-14.

Siegel, D. R., Esterly, J., Callanan, M. A., Wright, R., & Navarro, R. (2007). Conversations about science across activities in Mexican-descent families. *International Journal of Science Education, 29*(12), 1447-1466.

Snaddon, J. L., & Turner, E. C. (2007). A child's eye view of the insect world: perceptions of insect diversity. *Environmental Conservation, 34*(01), 33-35.

Snaddon, J. L., Turner, E. C., & Foster, W. A. (2008). Children's perceptions of rainforest biodiversity: Which animals have the lion's share of environmental awareness? *PLoS ONE, 3*(7), e2579.

Stoddart, T., Abrams, R., Gasper, E., & Canaday, D. (2000). Concept maps as assessment in science inquiry learning: A report of methodology. *International Journal of Science Education, 22*(12), 1221-1246.

Storksdieck, M., Ellenbogen, K., & Heimlich, J. E. (2005). Changing minds? Reassessing outcomes in free-choice environmental education. *Environmental Education Research, 11*(3), 353-369.

Strommen, E. (1995). Lions and tigers and bears, oh my! Children's conceptions of forests and their inhabitants. *Journal of Research in Science Teaching, 32*(7), 683-698.

Trowbridge, J. E., & Mintzes, J. (1988). Alternative conceptions in animal classification: A cross-age study. *Journal of Research in Science Teaching, 25*(7), 547-571.

Tunnicliffe, S. D. (2000). Conversations of family and primary school groups at robotic dinosaur exhibits in a museum: What do they talk about? *International Journal of Science Education 22*(7), 739-754.

Tunnicliffe, S. D., Lucas, A. M., & Osborne, J. (1997). School visits to zoos and museums: a missed educational opportunity? *International Journal of Science Education, 19*(9), 1039-1056.

Tunnicliffe, S. D., & Reiss, M. J. (1999). Building a model of the environment: How do children see animals? *Journal of Biological Education, 33*, 142-148.

Valle, A., & Callanan, M. A. (2006). Similarity comparisons and relational analogies in parent-child conversations about science topics. *Merrill-Palmer Quarterly, 52*(1), 96-124.

Van Zele, E., Lenaerts, J., & Wieme, W. (2004). Improving the usefulness of concept maps as a research tool for science education. *International Journal of Science Education, 26*(9), 1043-1064.

Venville, G. (2004). Young children learning about living things: A case study of conceptual change from ontological and social perspectives. *Journal of Research in Science Teaching, 41*(5), 449-480.

Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.

Yin, Y., & Shavelson, R. J. (2008). Application of generalizability theory to concept-map assessment research. *Applied Measurement in Education, 21*, 273-291.

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

Zimmerman, H., Reeve, S., & Bell, P. (2009). Family sense-making practices in science center conversations. *Science Education, 94*(3), 478-505.

BIOGRAPHICAL SKETCH

Jennifer Mesa earned a Bachelor of Science degree in Wildlife Ecology and Conservation and a Master of Education in Science Education from the University of Florida. After working as a middle school science teacher for three years in Gainesville, Florida, she returned to the University of Florida to work on a Doctorate of Philosophy in Curriculum and Instruction under the supervision of Dr. Linda Cronin-Jones. She is currently an elementary school science resource teacher in Gainesville, Florida.