University of Cincinnati

Date: 2/17/2014

I, Uzma Nooreen Maherally, hereby submit this original work as part of the requirements for the degree of Doctor of Education in Curriculum & Instruction.

It is entitled:
Development and Validation of the Life Sciences Assessment: A Measure of Preschool Children’s Conceptions of Basic Life Sciences

Student’s name: Uzma Nooreen Maherally

This work and its defense approved by:

Committee chair: Sally Moomaw, Ed.D.

Committee member: Mary Barbara Boat, Ph.D.

Committee member: Stephan Pelikan, Ph.D.
Development and Validation of the Life Sciences Assessment: A Measure of Preschool Children’s Conceptions of Basic Life Sciences

A dissertation submitted to the
Graduate School
of the University of Cincinnati
in partial fulfillment of the
requirements for the degree of

Doctor of Education

in the Department of Curriculum and Instruction:
Teaching and Learning of School Subjects
of the College of Education, Criminal Justice, and Human Services
by

Uzma Nooreen Maherally

M.Ed. Science Education, University of Southern Queensland, Australia
B.Sc. (Honors) Biology, University of Mauritius

February 2014

Committee Chair: Sally Moomaw, Ed.D.
Abstract

The purpose of this study was to develop and validate a science assessment tool termed the Life Sciences Assessment (LSA) in order to assess preschool children's conceptions of basic life sciences. The hypothesis was that the four sub-constructs, each of which can be measured through a series of questions on the LSA, will make a significant contribution to the latent construct, thereby causing the LSA to be a significant indicator of preschool children’s conceptions of life sciences. Two research questions pertained to this study “What is the contribution of each sub-construct to the latent construct on the LSA?” and “What is the technical adequacy of the LSA in measuring preschool children’s conceptions of basic life sciences?” The assessment tool, which consisted of a number of photographs and related questions, was tested with a convenience sample of 124 preschool children from four accredited child care centers in Cincinnati during the 2012-2013 school year. Confirmatory factor analysis was used to evaluate the hypothesized model. Results revealed a good model fit: \( \chi^2 (1, N = 124) = 0.929, p = .335; \) NFI = .992; CFI = 1.000; RMSEA = .000; ratio of the \( \chi^2 \) to the degrees of freedom = 0.929. In addition, all standardized regression coefficients demonstrated statistical significance at \( p < .001 \) (2-tailed) as well as practical significance (\( \beta > .3 \)), with moderate to high standardized regression coefficients ranging from .43 to .98. The reliability and validity assessments used to examine the technical adequacy of the instrument indicated that the LSA is a reliable and valid measure for this sample of preschool children: Cronbach’s alpha = .733; construct reliability = .714; content validity addressed in stages as required; construct validity established by CFA demonstrating statistical significance at \( \alpha = .001 \) (2-tailed) as well as practical significance; and successful minimization of potential threats to internal validity.
Dedication

To my husband, Mohammad Iqbal, for his unconditional love and constant support.

Thank you for always being by my side and for providing me with the strength to never give up during difficult times.
Acknowledgements

First and foremost, I would like to thank GOD for providing me with the blessings, privilege, courage, and perseverance to pursue this doctoral program and successfully complete it in spite of the many challenges faced during this remarkable journey.

There are several individuals to whom I would like to extend my thanks. I would like to start with my family. I thank my husband for his constant support and encouragement throughout my studies, and also for helping with data collection while simultaneously collecting data for his own study. I thank my father Iqbal Hussein and my mother Amina for stressing the importance of education since I was a little girl. I thank my father-in-law Aboo Twalib for always encouraging me to pursue my studies and my mother-in-law Hafez for always being available to talk to. I thank my brother-in-law Mohammad Nabiil for taking some photographs which I have used as part of my assessment tool. I thank my sister-in-law Dr. Zaynah Maherally for providing me with some feedback about the scoring system of my assessment tool.

I would also like to thank my dissertation committee. Dr. Sally Moomaw has been my doctoral advisor and her experience working with young children has contributed much to my knowledge about young children. As a committee member and director of Graduate Studies, Dr. Mary Boat was always available to provide her support and advice. I thank Dr. Stephan Pelikan, who took out time from his busy schedule to act as another committee member.

I would also like to thank some individuals who have directly or indirectly contributed to my dissertation. As a biology content expert with extensive classroom experience, Mrs. Lina Gunnoo provided her help in scoring categories of responses for a set of questions on the assessment tool. Dr. Wei Pan, a former committee member and professor for several of the statistics classes I was enrolled in, has made statistics understandable and exciting. His lectures
inspired me to undertake a research study involving multivariate statistics. I am grateful to Mohammad Irshaad Abdool, an Information Systems Engineering student, for his precious help with graphics and formatting issues. Dr. Holly Johnson, whom I see as a blessing in disguise during my doctoral studies, made life at UC seem more colorful. She is a careful listener and is always ready to provide her support. Finally, I greatly appreciate the assistance of the administrators, staff, teachers, parents, and children of the four child care centers where data collection took place.
Table of Contents

Abstract .......................................................................................................................... ii
Dedication ..................................................................................................................... iv
Acknowledgments ......................................................................................................... v
Chapter 1. Introduction ................................................................................................. 1
  Science and Young Children ...................................................................................... 1
  Misconceptions in Science ....................................................................................... 4
    Nature of Scientific Misconceptions ...................................................................... 4
    Reasons for the Occurrence of Scientific Misconceptions ................................. 6
    Relationship Between Language and Scientific Misconceptions .................... 7
Concept of Life and Living Things ............................................................................. 8
Theories Guiding this Study ....................................................................................... 10
  Folk Biology ............................................................................................................... 11
  Children’s Theory of Biology .................................................................................. 13
  Constructivist Theory .............................................................................................. 16
Children’s Conceptions of Plants, Animals, and Living and Non-Living Things .... 19
  Children’s Conceptions of Plants .......................................................................... 20
  Children’s Conceptions of Animals ....................................................................... 23
    Identification of Animals ...................................................................................... 24
    Reasons for Considering Items as Animals ......................................................... 26
Children’s Conceptions of Living and Non-Living Things ................................... 29
  Stages of the Understanding of the Concept of Life ............................................ 30
  Animism .................................................................................................................. 31
  Subsequent Research Findings Opposite to Piaget’s Findings ............................ 34
  Effect of Culture and Language on Conceptions of Living and Non-Living Thing ........................................................................................................................................ 37
Problem ................................................................................................................... 39
  Assessment in Education: Curriculum-Based Measurement ............................. 42
Significance .................................................................................................................. 46
Hypothesis ................................................................................................................... 50
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Questions</td>
<td>50</td>
</tr>
<tr>
<td>Chapter 2. Methods</td>
<td>51</td>
</tr>
<tr>
<td>Sampling</td>
<td>51</td>
</tr>
<tr>
<td>Participant Inclusion Criteria</td>
<td>51</td>
</tr>
<tr>
<td>Instrumentation</td>
<td>52</td>
</tr>
<tr>
<td>Instrument Design</td>
<td>53</td>
</tr>
<tr>
<td>Photographs Included in the LSA</td>
<td>53</td>
</tr>
<tr>
<td>Instrument Questions</td>
<td>54</td>
</tr>
<tr>
<td>Goals of the LSA</td>
<td>55</td>
</tr>
<tr>
<td>Measurement Scale</td>
<td>55</td>
</tr>
<tr>
<td>Pilot Testing and Revision</td>
<td>56</td>
</tr>
<tr>
<td>Instrument Administration</td>
<td>59</td>
</tr>
<tr>
<td>Procedure Prior to Instrument Administration</td>
<td>59</td>
</tr>
<tr>
<td>Procedure During Instrument Administration</td>
<td>59</td>
</tr>
<tr>
<td>Data Management</td>
<td>60</td>
</tr>
<tr>
<td>Data File Structure</td>
<td>60</td>
</tr>
<tr>
<td>Data Quality</td>
<td>60</td>
</tr>
<tr>
<td>Data Coding</td>
<td>60</td>
</tr>
<tr>
<td>Data Checking</td>
<td>61</td>
</tr>
<tr>
<td>Data Security</td>
<td>61</td>
</tr>
<tr>
<td>Data and Assumption Checks</td>
<td>61</td>
</tr>
<tr>
<td>Missing Data</td>
<td>61</td>
</tr>
<tr>
<td>Normality</td>
<td>62</td>
</tr>
<tr>
<td>Skewness and Kurtosis</td>
<td>62</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov (K-S) Test</td>
<td>63</td>
</tr>
<tr>
<td>Histograms</td>
<td>63</td>
</tr>
<tr>
<td>Linearity</td>
<td>66</td>
</tr>
<tr>
<td>Correlations Among Scales</td>
<td>66</td>
</tr>
<tr>
<td>Correlation Between Age in Months and Scores on the LSA .</td>
<td>66</td>
</tr>
<tr>
<td>Outliers</td>
<td>68</td>
</tr>
<tr>
<td>Univariate Outlier</td>
<td>69</td>
</tr>
</tbody>
</table>
Multivariate Outlier ......................................................... 69
Data Analysis Plan ............................................................... 69
Analysis of First Research Question ...................................... 70
Analysis of Second Research Question .................................... 71
Reliability Assessments ....................................................... 72
  Internal Consistency Reliability ........................................... 72
  Construct Reliability ....................................................... 73
  Reliability not Performed .................................................. 73
Validity Assessments .......................................................... 74
  Content Validity ............................................................. 74
  Construct Validity ........................................................... 76
  Internal Validity ............................................................. 76
  Criterion Validity ........................................................... 77
Chapter 3. Results .................................................................. 79
  Research Question 1 .......................................................... 79
  Research Question 2 .......................................................... 80
    Reliability ........................................................................... 80
      Internal Consistency Reliability ......................................... 80
      Construct Reliability ....................................................... 80
    Validity ............................................................................... 80
      Content Validity ............................................................. 80
      Construct Validity ........................................................... 81
      Internal Validity ............................................................. 81
Chapter 4. Discussion ............................................................. 82
  Major Findings ..................................................................... 82
    First Research Question .................................................... 82
    Second Research Question .................................................. 84
      Internal Consistency Reliability ......................................... 84
      Construct Reliability ....................................................... 84
      Content Validity ............................................................. 84
      Construct Validity ........................................................... 85
Internal Validity ................................................................. 85
Other Findings ............................................................... 86
Limitations ................................................................. 89
Implications ................................................................. 90
Considerations for Future Research ................................. 91
Conclusion ................................................................. 92
References ........................................................................ 94
Appendix A: Life Sciences Assessment ............................ 118
Appendix B: Photographs Used in the LSA ......................... 131
Appendix C: Table of Specifications for the Life Sciences Assessment ................................. 140
Appendix D: Tables ........................................................ 141
Appendix E: Data Figures .................................................. 146
List of Tables

Table 1. Challenges to Assessing Young Children, along with a Comparison of Traditional versus Curriculum-Based Assessments ................................................................. 45
Table 2. Demographic Information of Participants for Main Study .......................... 52
Table 3. Evolution of Questions from First Pilot Test to Main Study ....................... 58
Table A1. Life Sciences Assessment ....................................................................... 118
Table A2. Assessment Guide .................................................................................. 126
Table A3. Scoring Guide ......................................................................................... 127
Table A4. Examples and Explanations for Particular Scores for Part (d) Questions ...... 129
Table A5. Brief Description of Characteristics of Living Things and Examples .......... 130
Table C1. Table of Specifications for the Life Sciences Assessment ......................... 140
Table D1. Descriptive Statistics for the Scales of the LSA and the Total Measure .... 141
Table D2. Kolmogorov-Smirnov Values for the Scales of the LSA and the Total Measure .. 142
Table D3. Correlations Among Scales ..................................................................... 143
Table D4. Correlations Between Age in Months and Scores on the SLA ................. 144
Table D5. Test for Univariate Outliers on the Scales of the LSA .............................. 145
List of Figures

Figure B1. Fish ................................................................. 131
Figure B2. Butterfly ........................................................ 132
Figure B3. Dog ................................................................. 133
Figure B4. Strawberry ...................................................... 134
Figure B5. Grass ............................................................... 135
Figure B6. Flower ............................................................ 136
Figure B7. Snowman ......................................................... 137
Figure B8. Airplane .......................................................... 138
Figure B9. Clouds ............................................................ 139
Figure E1. Histograms Representing Total Scores and Frequencies on the LSA .......... 146
Figure E2. Scatterplots Comparing Age in Months to Scores on the LSA ................. 147
Figure E3. Hypothesized CFA Model for the LSA .............................................. 148
Figure E4. Standardized Factor Loadings of the Scales of the LSA ......................... 149
Chapter 1: Introduction

Science knowledge and processes are major factors that enable human beings to understand the ways the universe works. Since people need to be aware of how the universe works, it is important for science to be part of school curricula. According to the Center for Education in Science and Technology (2008), science is important in schools as students are provided with abundant and valuable knowledge that promotes understanding of the way things work and reasons for working as such. By employing the knowledge gained from exposure to science, children and students are able to comprehend new concepts, make well-informed decisions, and follow new interests. As science also assists in providing evidence for facts that are found in common instructional media such as books and television, it is thus a way to boost comprehension as well as help children and teenagers to preserve that information (Center for Education in Science & Technology, 2008).

Science and Young Children

While many adults perceive science as being a discrete body of knowledge, for young children, science means discovering the daily world around them, and this is particularly what they are fascinated to do during the whole day and every day (Conezio & French, 2002). Six reasons regarding why small children should have access to science (Eshach & Fried, 2005) include:

1) It is natural for children to have a good time when observing and thinking about nature.

2) Students experience positive attitudes toward science when they are in contact with science.

3) Students’ comprehension of the scientific concepts that are studied formally is enhanced as a result of early contact with scientific situations.
4) Children impact their subsequent development of scientific concepts with the utilization of scientifically related language during their early years.

5) Children have the ability to comprehend scientific concepts as well as to reason from a scientific perspective.

6) Scientific thinking can efficiently be developed as a result of exposure to science.

   French and Woodring (2013) state some other reasons regarding why science should be emphasized during the early childhood years. These are:

1) Young children are ready biologically and are willing to learn about the daily world, a world in which they need to survive and make progress. This learning can be converted into a form of science learning through a curriculum that encompasses children’s encounters with daily phenomena.

2) Science is rich with general content and assists children in learning vocabulary and creating a rich knowledge base.

3) Emphasizing science supports children’s learning of mathematics, which forms part of science-oriented inquiries. For example, children do activities like measuring, quantifying, classifying, and comparing of objects.

4) Inquiry-based science provides a context for teaching literacy. For example, it provides opportunities for read-aloud activities and talk related to science investigation as well as models for writing.

5) Science provides the opportunity to acquire vocabulary and take part in science conversations. For example, science encourages discourse forms like description, explanation, and argumentation.
Since young children have basic competence in science prior to formal schooling, policy recommendations include the fact that science should be considered as a vital component of a comprehensive, high-quality preschool program rather than an extra (Brenneman, Stevenson-Boyd, & Frede, 2009). Even though some science-based early childhood programs that have been implemented on a small scale have been regarded positively by children, teachers, and parents, at present, there is no action for their propagation and implementation on a large scale (French & Woodring, 2013).

Science is considered as important as a content area in preschool because of the natural connection with young children’s way of processing experience as well as their inborn desire to know about the functioning of the everyday world (French, 2004). Preschool science is regarded as a major area of the cognition and general knowledge domain by several state readiness standards (Greenfield et al., 2009). As such, recommendations include that courses focusing on the teaching of science at the early childhood level be integrated into teacher certification requirements as well as teacher preparation programs, and that professional development provide support to teachers to learn to integrate language, literacy, and mathematics into science activities based on inquiry (French & Woodring, 2013). A thorough review of the content and process emphases of 29 national and state pre-kindergarten/kindergarten science standards as well as 10 early childhood curricula has been performed by Greenfield and colleagues (Greenfield et al., 2009). They found that most of the content items could be grouped into three broad content areas, where the highest percentage of all entries were related to the Life Sciences (42%), followed by Physical/Energy Sciences (31%), and lastly Earth/Space Sciences (27%). These percentages clearly show the importance of life sciences at the early childhood level. As pointed out by Venville (2004), the majority of science curricula comprise the topic of living
things in the very early years of schooling in order for learning and teaching to be built upon this foundational knowledge and then extended throughout elementary and secondary school.

In order to study children’s conceptions of basic life sciences, an understanding of the nature of misconceptions in science as well as the reasons for the occurrence of misconceptions in science need to be reviewed. Information related to the present study, that is, children’s misconceptions about animals, plants, the living/non-living concept, the theories guiding their conceptions of life sciences, as well as curriculum-based measurement, are provided. Information regarding the pilot study of the development and validation of the assessment tool to measure preschool children’s conceptions of basic life sciences is also provided.

**Misconceptions in Science**

Research in science education over the past three decades has provided evidence of students’ non-scientific ideas in several science concepts (Dikmenli, 2010). According to Babai, Sekal, and Stavy (2010), a significant amount of evidence has shown that the various alternative conceptions that pupils hold come from the majority of science areas. Historically, a major goal of the research into non-scientific ideas was to find out conceptual difficulties in specific fields of study as well as to have a good understanding as to how these ideas originated (Modell, Michael, & Wenderoth, 2005).

**Nature of scientific misconceptions.**

In science classes, knowledge construction starts based on what students know when they come to the class (Bulunuz, Jarrett, & Bulunuz, 2008). Various sources of information for children exist, for example, the society in which the children live, care givers, several artifacts, technological media (Russell, 1993), adult influences (Tunnicliffe, Gatt, Agius, & Pizzuto, 2008), and out-of-school experiences (Tunnicliffe & Reiss, 1999). However, the ideas that
students often bring to class do not always correspond with scientific facts (Akamca, Ellez, & Hamurcu, 2009). These ideas, which reflect inaccurate understanding of scientific phenomena (Türkmen & Usta, 2007), are formed from students’ everyday experiences (Jimoyiannis & Kommis, 2003), are developed before students are instructed formally and also during their beginning school years, and may continue even after appropriate instruction (Dikmenli, 2010). These misconceptions tend not to change when scientific concepts are taught to students, that is, they are to a certain degree resistant to change (Ekici, Ekici & Aydin, 2007; Özay & Öztas, 2003). Misconceptions in science that persist after formal instruction can be due to everyday experiences, improper understanding from classroom instruction, and incorrect concepts by teachers and textbooks (Mohapatra & Bhaduria, 2009; Yip, 1998). A point to consider is that misconceptions do not occur only in young children, but in high school students as well as college students (Kubiatko & Prokop, 2007), that is, at all levels of school (Türkmen & Usta, 2007).

The ways in which learners process new information are affected by the misconceptions they hold (Atasoy, Kucuk, & Akdeniz, 2011). As Vosniadou (2008) points out, students do not know that their beliefs hinder their learning and reasoning and do not understand that their beliefs are not real facts but simply hypotheses that have the possibility of being tested and falsified. She further adds that students are not aware of other people’s beliefs and different points of view.

Research has revealed that there are various complex sources that give rise to conceptual and reasoning difficulties, which are difficult to remedy (Modell et al., 2005). Mental models refer to representations of an object or an event and reflect somebody’s knowledge about a phenomenon, knowledge that can be similar to or different from scientifically accepted
knowledge (Tunnicliffe & Reiss, 1999). In the event that students’ mental models do not match scientifically accepted concepts or that students inappropriately apply the model, they display misconceptions and reach final inappropriate answers because they experience conceptual or reasoning difficulties that interfere with their understanding of the subject (Modell et al., 2005). Even though learners in a science classroom are exposed to scientifically correct mental models, many of them still tend to hold these misconceptions (Bulunuz et al., 2008; Türkmen & Usta, 2007), or they tend to return to these misconceptions at the end of their courses (Sewell, 2002).

**Reasons for the occurrence of scientific misconceptions.**

Occurrence of misconceptions is due to several reasons as reported by Türkmen and Usta (2007): mental functioning levels; cognitive achievement; informal learning; and absenteeism. A major reason is that the teaching and learning of science concepts do not match pupils’ mental functioning levels. In this case, pupils are unable to understand abstract knowledge. A second reason concerns the cognitive achievement and the intelligence quotient (IQ) status of the students. Third, informal learning, which occurs outside of the formal classroom situation, also creates misconceptions. Fourth, absenteeism of students is another key reason for misconceptions. It is difficult for students to catch up when learning a concept if they have not attended all phases of the learning cycle. Meyer (1993) stated that many students add inconsistencies into their knowledge without recognizing the illogical merging of old and new ideas. In addition to the sources of misconceptions reported by Türkmen and Usta, another source is the terminology utilized during instruction since use of conflicting terms tend to create confusion in students, resulting in the need for teachers to be careful about the terminology they use and those found in textbooks (Dikmenli, 2010).
Not all students hold the same misconceptions about a concept. Misconceptions that students have may also be context specific, suggesting that teachers need to identify the misconceptions that apply to the context of the class (Wescott & Cunningham, 2005). The authors list geographical areas, gender, age, religious background, and generation as examples of factors that contribute to the variations in misconceptions among students. Therefore, investigating children’s ideas is a valuable move toward teaching them (Tunnicliffe et al., 2008).

**Relationship between language and scientific misconceptions.**

The inherent interest that young children have in science is mirrored in their daily dialogues (Mantzicopoulos, Patrick, & Samarakunavagan, 2008). Even before starting formal schooling, young children engage in conversations during everyday activities and ask questions about science (Fleer & Cahill, 2001; Kallery & Psillos, 2001; Korpan, Bisanz, Bisanz, Boehme, & Lynch, 1997). Children’s capacity to clarify scientific concepts and debate scientific ideas may depend greatly on context and convenience in selecting words (Salleh, Venville, & Treagust, 2007).

Clerk and Rutherford (2000) state that while language confusion may be destructive for science learning, language issues at times are disguised in the form of scientific misconceptions. According to them, in the event a pupil provides wrong answers to questions, teachers frequently conclude that the pupil holds true scientific misconceptions whereas the problems may be related to meaning in language rather than related to mental concepts.

Vygotsky (1986) put forward the idea of the ‘zone of proximal development’ to designate the potential for cognitive development. This zone is the restricted area of deep investigation for which someone is cognitively prepared and is completely developed only in the event of social and interactive communication. Vygotsky proposed that thought development is driven by
language since it is through language that new information can be labeled and utilized for interaction (Boudourides, 1998). Vygotsky declared that the difficulty with scientific concepts lies in children’s language development and thus also stressed that regarding scientific thinking, the original verbal definition is of utmost importance. Furthermore, he believed that word development is tied with conceptual development and constitutes moving from “everyday” to “schooled” or “scientific” concepts (Wertsch, 1990). While everyday concepts are closely linked to the word or object name or condition being considered (Salleh et al., 2007), scientific concepts, on the contrary, are orderly, considered separately from the instant image produced by the word and can be controlled by the mind (Gallimore & Tharp, 1990).

**Concept of Life and Living Things**

A valuable way of knowing how well biology instruction is being performed is to investigate students’ comprehension of the elementary concepts and to provide particular consideration to students’ misunderstanding of these concepts (Yorek, Şahin, & Aydin, 2009). One concept regarded as an important component in biology education (Yorek et al., 2009) and that represents a chief topic in the biology curriculum (Babai et al., 2010) is the concept of life and living things. There has been extensive research regarding children’s conception of living things for decades (Babai et al., 2010) throughout the world (Chen & Ku, 1997). Research on this topic began with the early work of Piaget (1929) as well as Laurendeau and Pinard (1962) decades later.

It is challenging to precisely define certain concepts in science. For example, even though biology is normally termed as a life science discipline exploring living things, an accurate definition of the concept of life cannot be made (Yorek et al., 2009). It is challenging to propose an accurate scientific definition due to scientific uncertainty as well as moral, legal, and
theological facets of this concept (Franklin, 2004). However, since there are far more living things that have not yet been classified as compared to the number of established living things, in the case of the future of biological diversity, it is essential that biology teachers understand the significance of the system of living things for humans (Yorek et al., 2009). Furthermore, comprehension of the concept of living things is essential to possessing a biological theory (Carey, 1985; Siegal & Peterson, 1999) and a requirement to enable successful participation in every biological lesson and learning activity (Salleh et al., 2007). Therefore, an important step is to be aware of the research that has explored early biological reasoning in general (Keil, 2003).

A large body of work (Carey, 1985; Inagaki & Hatano, 2004; Keil, 1994) has revealed that, during the preschool years, young children establish a naïve understanding of biology. This naïve understanding includes the capacity to differentiate between living and non-living things (Gelman, 1990; Gelman & Opfer, 2002; Inagaki & Hatano, 1996; Opfer & Siegler, 2004; Poulin-Dubois & Heroux, 1994; Rakison & Poulin-Dubois, 2001), awareness of what is found inside living things (Gelman & Wellman, 1991; Gottfried & Gelman, 2005), and comprehension of growth, digestion, and reproduction (Backscheider, Shatz, & Gelman, 1993; Bernstein & Cowan, 1975; Hirschfeld, 1995; Keil, 1994; Rosengren, Gelman, Kalish, & McCormick, 1991).

Considerable research in cognitive development has examined the differences children recognize between living and non-living things. In spite of the amount and nature of existing work in this area, there are important gaps in present understanding (Jipson & Gelman, 2007). The authors note that what is questionable is the way children reason about items that seem to move across the boundary between animate and inanimate, as they are prone to come across various objects that test this boundary in their daily lives. Moreover, Erickson, Keil and Lockhart
(2010) have claimed that “the extent to which children’s understanding of the biological domain is systematic and coherent remains unclear” (p. 392).

In the earliest work on children’s ontological comprehension, children regularly stated that non-living items like bicycles and clouds were alive (Bullock, 1985; Carey, 1985; Laurendeau & Pinard, 1962; Piaget, 1929). As reported by Jipson and Gelman (2007), this result can be interpreted in various ways: First, children and adults have a common understanding of the concept of life, but children have partial knowledge about the world and thus think that some non-living entities are alive; and second, it is possible that children and adults hold a different meaning of the term ‘alive’ (Carey, 1985). Therefore, when children are asked to determine whether specific items are alive or not, it is not clear whether and on what basis they distinguish between living and non-living kinds (Jipson & Gelman, 2007).

Attempts to investigate children’s ontological comprehension have focused on children’s reasoning regarding whether entities have particular characteristics or not (Jipson & Gelman, 2007). The accumulating evidence that children distinguish between living and non-living kinds contrasts sharply to Piaget’s early understanding about children’s reasoning. According to Jipson and Gelman (2007), however, it is early to infer that children discriminate between living and non-living things in terms of distinct ontological types because of the following reasons: first, the items considered in the majority of research paradigms do not mirror the variety of children’s real worlds; second, consideration of the specific characteristics children consider when making ontological distinctions is partial; and third, a focus on children’s reasoning regarding biological properties may overestimate their capacity to make clear domain distinctions.

**Theories Guiding this Study**
Three theoretical frameworks inform this study. These include 1) folk biology, 2) children’s theory of biology, and 3) constructivist theory. The theories are described below.

**Folk biology.**

The cognitive study of the way people categorize and mentally analyze the organic world is known as folk biology (Atran, n.d.). As pointed out by Anggoro, Waxman, and Medin (2005) and Anggoro (2006), a significant amount of research has emphasized concepts and reasoning regarding entities in the biological world, with particular attention to folk biologic knowledge, that is, the individual’s daily instinctive knowledge about living things. Anggoro provides the following information regarding folk biology. People’s reasoning about biological events is based on instinctive theories that assist them in moving along in the natural world. Folk biologic knowledge is a type of folk theory that is differentiated from formal scientific teachings. This differentiation is due to the fact that it might not conform to formal standards relative to western or formal science. She provides two examples to illustrate this type of knowledge: first, in the case of corals, these are considered as living organisms according to western science teachings but are simultaneously considered as inanimate items from the point of view of a community’s folk theory; and second, in the case of non-living natural objects like rocks, these are considered as non-living things according to western science teachings but regarded as living things in another community. Therefore, she concludes that folk biology appears to influence categorization of objects in the natural world and is also present in cognitive processes like causal reasoning and induction. Seen in this light, it can be said that the concepts underlying natural phenomena are often regarded differently by children and scientists (Carey, 1985). Many young children are unable to tell the difference between living and nonliving things in ways that are considered scientifically acceptable (Venville, 2004). It is challenging for children even at 10
years old to understand the scope of “living thing” (Hatano et al., 1993). As Anggoro et al. (2005) state, folk biologic concepts such as the living thing concept are difficult to attain, thus partly reflecting the difficulty children experience in deducing relationships among these concepts.

With the huge amount of research that has emphasized “folk-biological” knowledge, there is ample general agreement across distinct measures, distinct lab groups, and different decades of research, that recognition of the comprehensive concept “living thing,” one that comprises members of both the animal and plant kingdoms, is a fairly late and difficult developmental accomplishment (Anggoro, 2006; Hatano & Inagaki, 1994). According to Anggoro (2006), folk biology arises early in development, but the development of a mature folk biological knowledge is a challenging process. As stated by Anggoro, Waxman, and Medin (2008), at least some primary folk-biological concepts (such as “animal”) appear early, and considerably before animal names. Infants usually show interest in animate items and are attracted by animate objects, including faces, eyes, and voluntary, biological movement (Bertenthal, 1993; Johnson, Slaughter, & Carey, 1998; Poulin-Dubois & Shultz, 1990). They can differentiate between animate and inanimate items (Woodward, Sommerville, & Guajardo, 2001). Furthermore, Piaget (1954) had much earlier discussed the fact that children have a rudimentary belief, as evidenced by their inclination to incorrectly assign animacy to inanimate items like clouds and bicycles that seem to move by themselves or display goal-directed behavior. As a result of the discovery of childhood animism, Piaget declared that children have a very distinct comprehension of critical folk-biologic concepts like “animal” and “living thing,” and have up until now not figured out the relation between them. Maybe not surprisingly, based on this, ideas of animacy are apparent in preschool children’s reasoning (Gelman, 1990).
Children’s theory of biology.

Children’s theory of biology has been summarized by Venville (2004) as follows. As children develop an understanding of the concept of life, the links between concepts help to shape the children’s developing theory of biology. Children who are able to establish new concepts and beliefs corresponding to core biological theory are those who can understand the concept of life in ways that biologists do. Even though the ability to distinguish between living and non-living things is of utmost importance for the beginning of a child’s theory of biology and for biological thought (Siegal & Peterson, 1999; Slaughter, Jaakkola, & Carey, 1999), Venville acknowledged that a great percentage of students in late elementary school do not have a sound understanding of the concept of living corresponding to biological theory. For example, children younger than 10 years old confuse biological phenomena with psychological phenomena; they do not comprehend biological phenomena as “biological” at all, but rather misunderstand them as “psychological” and thus have a psychology/biology theory that cannot be distinguished from each other (Carey, 1985). However, theoretical and empirical investigations demonstrate evidence that children have biological knowledge at rather younger ages, as Carey (1985) had suggested. Along this line, Hatano and Inagaki (1994), Inagaki (1990), and Inagaki and Hatano (1993) showed that young children’s biological knowledge is considerably affected by early experiences with live organisms (Jaakkola & Slaughter, 2002; Teixeira, 2000). In the case of preschool children, it should be noted that proof of preschool children’s naïve biology has been gathered from a huge amount of research, exhaustive in its methodological precision but narrow in its unifying scope (Margett & Witherington, 2011). Individual studies regularly focus on some subclasses of the living/non-living kinds distinction, such as comparing plants with animals or comparing animals with artifacts, or emphasize some
biological processes (for example, grow or regrow) to the exclusion of others (Erickson et al., 2010).

Review of research by Margett and Witherington (2011) indicates two different theoretical accounts that presently communicate the idea of the development of children’s view of living kinds. While one account suggests that a qualitative reorganization occurs during childhood in the exact structure of children’s conceptualization of living kinds (Carey, 1985), the alternate account suggests a core competence notion of children’s conceptualization of living kinds (Gelman, 2003; Gelman & Wellman, 1991; Keil, 1994; Keil, Levin, Richman, & Guttheil, 1999).

Regarding the first account, as children mature, there is an improvement in thinking regarding the state of living. According to Carey (1985), the qualitative reorganization constitutes an essential progression in thinking from preschool children’s wrong dependence on naïve psychology to the school-aged child’s appropriate dependence on naïve biology. Preschool children conceive life in exclusively psychological rather than extensively biological terms as they merge the state of being living with psychological agency. This conceptualization is subjected to extensive reorganization only by late childhood, and it sufficiently mirrors the biologically oriented living-kinds understanding of the adult.

Regarding the second account, the scholars (Gelman, 2003; Gelman & Wellman, 1991; Keil, 1994; Keil et al., 1999) maintain that the elementary biological framework for distinguishing between living and non-living kinds is already apparent in the preschool period, and although not prominent, gives the young child a theoretical series of regularities for organizing the living world (Erickson et al., 2010). According to Margett and Witherington (2011), it is this approach that constitutes advances in conceptual development regarding the
living-kinds concept, rather than key shifts in the structural organization of this concept, that is, there is a continuous amplification of an already existing competency, extensive in scope but restricted in accuracy and detail.

As Margett and Witherington (2011) explain, there are two ways by which researchers view this process of amplification: first, dependence on experience with biological items, and second, dependence on language development. Regarding experience with biological items, depending on situations, it is plausible that young children will reason about biological items in terms of humans due to their greater knowledge about humans as compared to the other items (Inagaki & Hatano, 2006). A higher degree of familiarity with a broad collection of living things encourages higher degree of proficiency in young children’s use of their core naïve biology (Margett & Witherington, 2011). This is exemplified in some studies (Atran et al., 2001; Medin & Atran, 2004) that investigated children who grew up in urban versus rural, less industrialized regions.

According to other researchers (Greif, Kemler Nelson, Keil, & Gutierrez, 2006; Waxman, 2005), the process of amplification depends to a great extent on language development, with preschool children developing greater skill in linguistically matching conceptual abstractions (Margett & Witherington, 2011). Anggoro (2006) suggests that language affects the development of folk biology, especially the naming practices of the community in which children grow. Investigations using child-introduced questions to evaluate young children’s naïve biology, that is, a way for the children to utilize their own words instead of adapting to the researcher’s terminology, propose that preschool children do display conceptual awareness of living kinds (Greif et al., 2006; Kemler Nelson, Egan, & Holt, 2001). Therefore, deficiency in achievement displayed when young children are asked to clearly differentiate items as alive or living arise
from matching challenges connected with linking the term alive with their living-kinds conceptualization, rather than from conceptual knowledge (Margett & Witherington, 2011). To summarize, regardless of the particular experiential actions involved, the core competence approach calls upon processes of aptitude improvement and expansion of application rather than processes of organizational restructuring in revealing preschool children’s conceptual development (Margett & Witherington, 2011).

**Constructivist theory.**

As a psychological theory, constructivism originates from the expanding field of cognitive sciences (Fosnot, 2005) and is a major theory used in science education in terms of curriculum design and research (Ferguson, 2007). Instructional methods that are based on the constructivist theory, that is, methods that direct students to acquire information by themselves instead of being presented with the information, have provided education researchers with a favorite teaching technique for many decades (Sweller, 2009).

Two famous scholars involved in the development of constructivist theories include Jean Piaget and Lev Vygotsky. Their theories are described by Ozer (2004). Piaget explains the learning process in the following way:

- by schemes (mental organization of information on the way things work)
- by assimilation (putting new information into schemes)
- by accommodation (completely changing existing schemes or generating new ones)

Discovery is the basis of Piaget’s developmental theory of learning and constructivism. The motivation for learning arises from the willingness of the learner to conform to his environment, and thus to establish a state of equilibrium between schemes and the environment.
New learning takes place due to prolonged encounters among existing schemes, assimilation, accommodation, and equilibrium.

According to Vygotsky (1986), constructivism is centered on a social aspect due to the emphasis of culture and social context. Vygotsky puts forward that learning occurs in collaboration with development, and children undergo cognitive development in the context of socialization and education. For him, the concept of a zone of proximal development (the difference between what a learner can do without help and what he/she can do with help) acknowledges that cognitive development is restrained to a certain range at a definite age. However, social encounters, in the form of, for example, help from a mentor, allow students to understand concepts and schemes that they do not know on their own. Essential cognitive tools affiliated with culture, such as history, social context, traditions, language, and religion, completely change children’s perceptual, attention, and memory abilities. In order for learning to result, children originally get acquainted with the social environment on an interpersonal level, after which they internalize this encounter. The previous ideas as well as the new encounters have an impact on the children, who then build new ideas.

Constructivism is different from behaviorism, which takes into account external reinforcements and direct instruction, where students are given directions about what to do (Schwartz, Lindgren, & Lewis, 2009). These authors acknowledge that direct instruction has the possibility of being an effective way of learning, but only when learners have enough prior knowledge to build new understanding based on the guidelines they are given about what to do. For these authors, constructivism is a broad facet of learning and a means of teaching that allows understanding of pupils’ capacities to build new knowledge at a time when teaching is not being
carried out. They stress that constructivist teaching outcomes should benefit students in their learning even in the absence of the teacher.

Instead of considering learning to be a product of teaching, the constructivist learning theory takes into account what students do when they obtain new information (Sewell, 2002). It is a theory of learning that assumes that children build knowledge based on the connection between their own ideas and that of their encounters in the social and physical world (Chaillé, 2008). For a long time, constructivism has been regarded as a theory of learning where learners construct mental representations by interacting in suitable types of active cognitive processing while learning (Mayer, 2009). The theory puts forward that learners have different experiences, knowledge and ideas, with some of these ideas being in contrast to what is scientifically acceptable (Colburn, 2000). This means that people perceive similar things in different ways based on their experiences (Shapiro, 2008) because they build mental representations of the outside world that they can utilize in order to function in that world (Sweller, 2009). This point is extended by Mayer (2009), who states that engagement in cognitive processes during learning enables active construction of mental representations in learners’ working memories. Moreover, active learning takes place when learners participate in proper cognitive processing during learning, which in turn results in the creation of cognitive representations (Mayer, 2001; 2008).

In accordance with the constructivist learning theory, teaching should contribute to proper cognitive processing during learning (Mayer, 2001, 2008).

Because proponents of constructivism believe that existing knowledge influences the learning process (Demirci, 2009), all learning is considered to be related to constructivism (Sweller, 2009). According to Sweller (2009), there is no theorist who disagrees with this depiction of learning. Furthermore, learning from the constructivist perspective is a self-
regulatory process, as when learners find new knowledge in conflict with existing knowledge and attempt to make meaning of the world (Fosnot, 2005; Ferguson, 2007). Because children are always attempting to make sense of the world and develop their ideas, motivation for learning is considered as natural from the constructivist viewpoint (Chaillé, 2008).

Thus, constructivist learning theory in science is a way for students to modify their beliefs to make them more scientific (Colburn 2000). On the whole, constructivists recognize that in order for meaningful learning to occur, learners need to build new knowledge actively (Demirci, 2009). This is because traditional approaches, which focus on memorization of facts and recitation, result in superficial understanding of scientific concepts (Türkmen & Usta, 2007).

As misconceptions and conflict work together to promote meaningful learning, presenting learners with information that is different to their beliefs causes them to examine the contrast between their current knowledge with the newly presented knowledge (Sewell, 2002). Following this, students have four choices: removing the current knowledge (wrong beliefs), changing the current knowledge to fit the new knowledge, changing the new knowledge to fit the current knowledge, or discarding the information presented (Sewell, 2002). Students then decide whether they will accept the new knowledge, which they often do, but only when they are unable to understand the world using their current knowledge (Sewell, 2002). This is in line with what radical constructivists say about learning: that “it is ultimately the individual who does the learning” (Bodner, 2004, p. 621). The process whereby individuals try to understand and find logical explanations for their experiences is often examined in studies dependent on the constructivist theory, for example, studies on meaning-making, building knowledge, understanding different views, and conceptual change (Ferguson, 2007).

Children’s Conceptions of Plants, Animals, and Living and Non-living Things
In order to inform design of the instrument, a literature review was conducted specific to children’s understanding of living things. Background regarding children’s conceptions of plants, animals, as well as living and non-living things, is provided below and separated into three themes.

**Children’s conceptions of plants.**

Plants are major members of the environment. They are a chief part of the scenery in the life of children, who come across them in various forms: potted plants, cut flowers, vegetables, and fruits as well as in images in the form of pictures, fabrics, and illustrations in books and greeting cards (Tunnicliffe, 2000).

Previous research regarding children’s conceptions of plants has reached the conclusion that children have limited knowledge about plants. They use their sensory and personal interactions with the environment to build their ideas about natural events (Tunnicliffe, 2001) and do not believe that plants are living things (Carey, 1985; Tunnicliffe, 2001). When they cannot attribute a name to an unfamiliar specimen, they refer to it as a plant, but usually consider mostly flowering plants as plants (Bell, 1981a).

Some studies have investigated pupils’ ideas about plants and have reported that while categorizing plants, pupils valued them mainly with respect to their anatomic features and outlook. Research carried out in New Zealand revealed that pupils between the ages of 13 and 15 held a more limited meaning of the word “plant” as compared to the one used in science, and they failed to view trees as plants (Bell, 1981a). Some pupils believed that non-living examples or other organisms were plants only due to their possession of structures that looked like a flower or stem, implying that they did not regard an organism as a plant unless it had particular parts.
like a flower or stem. Rymell (1999) found that children thought of plants as having no trunk and growing on the ground.

In the case of Maltese 4- and 5-year-old children, when asked to name plants, about 50% of the children mentioned flowers, 40% mentioned trees, and the rest simply repeated “plants” (Gatt, Tunnicliffe, Borg, & Lautier, 2007). In some cases, the children considered superordinate categories along with names of particular species; for example, one child believed that flowers and daisies were at the same level. Children had more things to say about flowers and trees than about the word “plant.” The most often cited plants included the rose, the apple tree, and the orange tree, followed by the sunflower. Thorns and nettles, as well as cacti, were mentioned by some children. In general, the authors indicated that the children’s main mental model consisted of a plant as being something small and green, and possessing a stalk and leaves, thereby limiting the concept of plant to small herbaceous angiosperms. Instead of considering flowers, fruits, and trees as subgroups of plants, the children believed they formed a separate group from plants. In many cases, they used only one characteristic to categorize an item as a plant; for example, a lettuce was a plant because it was green, but a cactus was not a plant because it had spines. Some children focused on plant parts instead of the whole item, for example, by noting the color of the flower first. They also had trouble differentiating between the names of plant parts and the plants themselves, for example, confusing petals with the plant’s name. The children’s beliefs only reflected everyday understanding of botany rather than the concept of plant kingdom in science.

In another study conducted in the United States, Barman, Stein, McNair, and Barman (2006) investigated K-8 students’ ideas about plants by showing them pictures of plants and pictures of objects that are not plants. Through an interview, the students were required to identify which pictures depicted plants and to explain the common feature of all plants. The
investigators found that over 95% of students succeeded in identifying a flower, a bush, and a fern as plants. While 75% of students recognized that the bread mold was not a plant, about half of the students thought that the mushroom was a plant. Interestingly, more K-5 students knew that a seed was a plant as compared to students in grades 6-8. Furthermore, more grades 3-5 students determined that grass, oak tree, and pine tree are plants as compared to students in grades 6 to 8. The authors further revealed that students associated the possession of leaves and stem, the green color, and growing in the ground as characteristics of plants. Considering the color green as a characteristic of plants was also found by Gatt et. al (2007). Considering growth on the ground as a characteristic of plants was also found by Rymell (1999). Barman and colleagues also found that students held misconceptions due to limited classification skills and not content knowledge; for example, when asked if a tree is a plant, a student replied negatively and pointed out that it is a tree. Another student pointed out that a seed is not a plant due to the absence of blooms or flowers. In general, the seeds and mushroom proved the most challenging to categorize. Failing to view trees as plants and considering an organism to be a plant due to the possession of flowers was found in Bell’s (1981a) study too.

The findings of Gatt et al. (2007) and Barman et al. (2006) regarding the consideration of some items as separate groups from plants rather than subgroups of plants, had previously been found in Chen and Ku’s (1997) study. In this study, in regards to naming “living things” and “things of life” by 2nd, 4th, and 6th grade Taiwanese aboriginal students, it was deduced that the term “plant” was considered as a class on the same level as the examples children gave, that is, tree, flower, and grass.

Wandersee and Schussler have (1999) debated the reasons for people’s greater interest in animals than in plants. They introduced the term plant blindness, that is, failure to see or notice
plants in the environment. According to them, plant blindness results in the incapacity to recognize the importance of plants for the atmosphere and for human life as well as to acknowledge the special aesthetics and biological characteristics in the plant kingdom. Additionally, they stated that it brings on misleading, anthropocentric judgment of plants as inferior to animals. A point to note is that teachers are responsible for causing students to hold misconceptions about plants when they spend more instructional time on the animal kingdom than on the plant kingdom (Patrick & Tunnicliffe, 2011).

**Children’s conceptions of animals.**

Living organisms have an important place in children's lives. Children learn about animals and plants from their earliest moments (Keil, 1979), although animals form the highest proportion of words in a child's first vocabulary (Rinsland, 1946). Regarding animals, children are exposed to animated cartoons, fictional stories, books (Barman, Barman, Cox, Kay, & Goldston, 2000), as well as live animals. The knowledge they acquire about animals is affected by the amount of knowledge they already have, environmental and sociological influences, as well as sources of knowledge that contribute to their understanding (Tunnicliffe et al., 2008).

The characteristics that children use to provide descriptions of animals do not reflect characteristics pertaining to all animals (Bell, 1981b). Since children’s understanding of the concept animal is more restricted than that used in science, the confusion that children experience regarding the scientific meaning and the common meaning of the word “animal” contributes to children’s misconceptions, denoting that difficulties arise in students when they are confronted with the conventional concept attributed to the word animal and the scientific meaning (Tunnicliffe et al., 2008).
Students’ understanding of animals can be promoted by the textbooks and science trade books utilized in the classroom. Review of a typical sample of these books has aroused some concerns (Barman et al., 2000): first, the majority of trade books did not utilize the word animal except when considering vertebrates, and more specifically mammals; second, books relating to insects or snakes frequently utilized the word *critter* or *creature* rather than the word animal; and third, textbooks meant for elementary students did not consist of discussion about what makes an organism an animal. Barman and colleagues also found a deficiency of any definitions of *animal* in the trade books and the exclusion of this information in the elementary science textbooks. They further noticed that it is only starting at the middle level that textbooks determine particular features that make an organism an animal. In addition, only two out of three middle level textbooks examined pinpointed human beings as animals.

Various studies have employed methodologies to inquire into the ideas of children of different age groups about animals: through interviews, by showing pictures or flashcards of items, and through preserved specimens. In general, in order to know the children’s ideas, they were asked to name animals, to name the items shown to them, to classify items shown to them as animals or non-animals, and to mention the features that caused an item to be an animal. Their responses were categorized in two themes as described in the following sections: identification of animals and reasons for considering items as animals.

**Identification of animals.**

In studies investigating the identification of animals, students had correct conceptions of animals in some instances while in other instances their responses did not reflect scientific concepts. In the replication of Bell’s (1981b) New Zealand study in Barcelona, Spain with 8- to 16-year-old students, Villalbi and Lucas (1991) used a set of 19 cards illustrating seagull, cow,
spider, worm, herb, cat, mushroom, sardine, child, frog, snail, elephant, snake, fire, lioness, whale, car, tree, and butterfly. They reported that contrary to Bell’s study, most of the children in their sample could correctly classify animals. However, similar to Bell’s study, their study sample yielded classification of spider, worm, and butterfly as non-animals by some students. In general, students gave a greater number of correct than incorrect answers. Furthermore, their study demonstrated that about one tenth of the participants considered birds, fish, and reptiles as comparable sets rather than subsets of animals. As for the card depicting the child, 91% of the younger children believed it was not an animal, followed by 65% of the 12-year-olds and 20% of the 15-year-olds.

In a study with students from second, fourth, and sixth grades (Chen & Ku, 1998), students were shown 12 flashcards including animal items such as tiger, snail, fish, girl, beetle, and snake, and non-living items such as car, boat, alarm clock, cloud, lightening, and sun. Human beings seemed to be the most problematic for the students to identify as an animal. However, they correctly identified that man-made and natural items were not animals, even though they could move.

In a national study regarding K-8 students’ ideas about what organisms they thought were animals (Barman et al., 2000), the researchers showed pictures of 18 organisms, namely, dog, snail, bird, girl, bear, flowering plant, jellyfish, frog, mushroom, lizard, butterfly, oak tree, fish, spider, earthworm, elephant, snake, and octopus. They found that a certain percentage of students either lacked confidence that humans are animals or did not regard humans as animals: 80% of primary students, 68% of students of grades 3 to 5, and nearly 50% of middle school students. Most students were right about the flower, mushroom, and oak tree as being non-animals. Most K-8 students rightly determined that the dog, bear, elephant, jellyfish, and octopus were animals.
Interestingly, K-2 students rightly identified a greater variety of organisms, such as bird, frog, lizard, butterfly, fish, spider, earthworm, and snake, as animals as compared to the older students. Another interesting finding was that a greater number of grades 3-8 students found it more difficult than K-2 students in identifying some vertebrates and invertebrates as animals.

In a more recent study, Tunnicliffe et al. (2008) explored 4-and 5-year-old Maltese children’s knowledge on the range of animals they knew and also showed them 11 photo cards depicting horse, sheep, dog, cat, elephant, dolphin, zebra, ladybird, spider, pig, and mouse. In addition to the animals depicted on the photocards, the children mentioned various other animals, with the most commonly cited ones being tiger and lion. More than 50% of the older children stated that the spider, dolphin, and ladybird were not animals. The only item which all children could identify as an animal was the horse. About two thirds of the younger children classified the dolphin as a fish and not as an animal. The authors further pointed out that the majority of the younger children viewed animals as mainly large and terrestrial, such as animals kept at home as pets, farm animals, and jungle animals.

**Reasons for considering items as animals.**

Research has shown that students have provided various reasons for why they considered an item as an animal, with the reasons being either scientifically acceptable or not. The main reasons provided by participants included habitat, behavior, appearance, reproduction, movement, growth, and breathing. These reasons, as well as others obtained in individual studies, are provided below.

Chen and Ku (1998) reported that fewer than 50% of the definitions of animals provided by the second, fourth, and sixth graders who participated in their study, reflected scientifically accepted attributes of animals. The most often cited attributes included eating and movement, as
well as features of animals such as appendages, habitats, behavior, and external appearances. The ability to attack was also mentioned. The majority of children provided reasons that made it clear that they depended on visual features of items to explain why they considered an item to be an animal. In some instances, there was mention of living organism attributes limited to life, reproduction, breathing, and growing up, as well as anthropocentric attributes (related to humans and human feelings). In the case of human beings, reasons pertaining to living organism, human behavior, lifestyle, appearance, evolution, or religion were cited. Even though students could not provide non-animal attributes for why man-made objects were not animals, they correctly provided reasons related to individual objects’ characteristics to prove these were not animals.

The dependence on visual features was confirmed by Tunnicliffe and Reiss (1999), who investigated children’s conceptions of six preserved specimens of animals, namely, armadillo, stoat, crab, starling, gecko, and stag beetle. The children were 5-, 8-, 10-, and 14-year-old pupils of varying abilities in the United Kingdom. The authors stated that pupils named the animals mostly according to their anatomy (87%), followed by behavior (10%), or habitat (3%). In some cases, pupils associated the anatomical features to the animals’ habitats and behaviors. According to Tunnicliffe and Reiss, young children usually identified striking characteristics of animals when they saw the live organisms, and these characteristics gather in children’s mental models of varying types of animals. As compared to older children, who group animals according to taxonomy, habitat, behavior, and anatomical structures, the authors found that young children referred to only anatomical structures to group animals.

In another study (Barman et al., 2000), the majority of reasons provided by students for why an organism is an animal included movement, having body parts and fur, being a pet, catching/eating food, living in water or a house, breathing, being big, reproduction, caring and
nursing for young ones, and being multicellular. The most common reasons for why an organism is not an animal included the inability to perform the following: move, breathe, catch food, reproduce, and nurse and care for the young ones, absence of fur and some body parts, being a person, being too small, being an insect, and not being a pet. Of particular interest is that when K-5 students provided criteria for why an organism can be an animal, they failed to provide the opposing criteria for why an organism cannot be an animal.

In a more recent study (Tunnicliffe et al., 2008), the criteria which children employed to justify that an item was an animal included appearance (such as number of legs, or presence of tail or fur), size, noise production, habitat, and personal experience with a particular animal, among other questionable answers. No children in the sample attributed scientific features of living things, such as respiration, growth, and reproduction, in justifying their decision about an item being an animal or not.

It appears that, irrespective of age, pupils commonly utilize the word animal in a way that is analogous to the biological concept of mammal (Bell, 1981b; Bell & Freyberg, 1985; Towbridge & Mintzes, 1988). For example, in Chen and Ku’s (1997) study, in regards to the naming of animals as “living things” and “things of life” by second, fourth, and sixth grade Taiwanese aboriginal students, it was deduced that children thought the term “animal” was a class on the same level as human and some other animals like lion, tiger, bird, cat, and dog. As Salleh et al. (2007) put forward, young children are inclined to consider items such as insects, fish, and humans as different from their concept of animal. It has also been found that children who are bilingual have a wider notion of the concept of animal as opposed to children who studied in other English speaking countries (Tema, 1989; Villalbi & Lucas, 1991). These two studies revealed that, when interviewed in their native language, the confusion between every
day and scientific meanings of the word animal was less evident for the bilingual children in Barcelona and South Africa. These pupils could distinguish animals from plants not only on the basis of movement, but also based on habitat, external appearance, and body functions. To explain the misunderstanding regarding the concept of animal, one can refer to Vygotsky’s (1986) statement that the difficulty students experience is not due to the complexity of the scientific meaning, but is rather due to the confusion between the conventional and scientific understandings of children. As pointed out by Villalbi and Lucas (1991), it is thus possible that there are facets of the English language that cause the learning of this concept, as well as connected scientific concepts, to be challenging.

**Children’s conceptions of living and non-living things.**

The National Science Content Standards in the United States describe what students should know, understand, and be capable of doing in the natural sciences during K-12 education (National Research Council, 1996). Life science is one of the eight categories of the content standards. Most science curricula include topics on living things in the early years of schooling so that foundational knowledge is available for teaching and learning to expand upon in elementary and secondary school. Even though living organisms are essential components of the environment and occupy an important place in children’s lives, there is little understanding as to how children view them (Tunnicliffe, 2000).

An important way in which objects may differ from one another is in their possession of life. Differentiating between living and non-living things is a principal part of biology, various branches of philosophy, as well as children’s attempts to make sense of their environment; therefore, being aware that an object possesses life enables children to deduce several other
substantial facts about it: for example, ability to grow, ability to take in nutrients, ability to breathe, ability to reproduce, etc. (Hatano et al., 1993; Richards & Siegler, 1986).

*Stages of the understanding of the concept of life.*

The majority of research that has emphasized children’s comprehension of the scientific concept of living has expanded from original interviews of Piaget (1929), who studied the concepts of consciousness and animism that are closely related to the idea of living and non-living things (Salleh et al., 2007). Of particular concern to this study is the concept of animism. As Salleh et al. (2007) put forward, assigning human or animal features (like personality or capacity to follow) to inanimate items, such as dolls and the moon, is referred to as animism. In Piaget’s (1930, 1951) early work, he stated that the child has an essentially animistic attitude toward all events since action is comprehended egocentrically on the basis of human behavior. The emphasis here is on learning to differentiate between animate and inanimate items. Piaget indicated stages of development for the concept of life. He designated a particular order to describe the change from a completely animistic notion to more differentiated concepts. Since he regarded the understanding of the concept of life as very important, he proposed four stages of understanding through which children advance. Children up to 6 years of age are in stage one and ascribe life to things that show some activity (for example, a ticking clock) or usefulness (for example, a spoon to eat with). These things comprise people, other animals, vehicles, the sun, and ovens, but exclude trees and other plants. Children aged 6 to 8 are in stage two and attribute life only to things that show some kind of movement, thus narrowing their definition. In this case, for example, people, other animals, vehicles, and the sun are regarded as alive while items like ovens, trees, and other plants are not. Children aged 8 to 12 are in stage three and consider only things that move independently as alive. This set of things includes, for example, people,
other animals, and the sun. Lastly, children aged 12 and higher are in stage four and ascribe life to either people and other animals only, or to people, other animals, and plants only. The child accepts the more precise criteria disconnecting biological from inanimate systems in late childhood (Bullock, 1985). Piaget inferred this to a great extent by analyzing children’s explanations regarding naturally occurring phenomena, such as the movement of water and clouds, and regarding the workings of machines, like steam engines and bicycles (Bullock, 1985).

Animism.

Piaget (1929) was the first researcher who brought up the idea that life was associated with motion because of an animistic thinking during early childhood. He took into account several familiar items such as plants, animals, vehicles, non-moving terrestrial objects, and involuntarily moving extraterrestrial objects like the moon, wind, and lightning, and asked the question “Is object X alive?” to the children followed by asking for explanation for their decisions. He first demonstrated that children are animistic by presenting the fact that 12-year-old children tended to disagree that plants are alive, but assigned life status to some non-living items, particularly those that seemed to move independently (see also Klingberg, 1957; Klingensmith, 1953; Laurendeau & Pinard, 1962; Russell & Dennis, 1939). The hardship in setting up the scope of the concept has been supported more recently in a variety of populations (e.g., Anggoro et al., 2008; Babai et al., 2010; Carey, 1985; Hatano, et al., 1993; Leddon, Waxman, & Medin, 2008; Margett & Witherington, 2011; Opfer & Siegler, 2004; Richards & Siegler, 1984; Stavy & Wax, 1989; Yorek et al., 2009). For example, in Hatano et al.’s (1993) study, the percentage of Israeli children from second and fourth grades who were prone to classify plants as non-living objects amounted to over 40%. Another example is that of Anggoro
et al.’s (2008) investigation with 4-year-old, 6-year-old, and 9-year-old American and Indonesian children, where some children from all the three age groups sorted humans and non-human animals together and excluded plants and other items when probed with the predicate “alive.” A third example is that of Yorek et al.’s (2009) study with ninth graders. In regards to naming 10 living things, the most frequently mentioned living things were animals, ranked in the order of human, dog, cat, and bird. The proportions of animals and plants that students mentioned were 80% and 13.4% respectively, with the rest being attributed to other living things. For students who mentioned at least one plant name, the authors found that the plant was ranked sixth in the list of 10 living things. They further concluded that cognitive construction of the life concept is first associated with animals, particularly humans, followed by plant and then other living things. Similar findings were obtained in Babai et al.’s (2010) and Margett and Witherington’s (2011) studies with 10th graders and preschoolers respectively. The rate of accurate classifications of grayscale drawings of animals, plants, inanimate static objects, and inanimate dynamic objects into the category of living objects, was significantly higher for animals than for plants (Babai et al., 2010). Preschoolers presented with real-time video clips of objects including plants, animals, mobile artifacts, and immobile artifacts tended to classify animals as living to a greater extent than plants (Margett & Witherington, 2011). In general, these investigations have emphasized that children correctly assigned life status to animate items, and were frequently capable of noting that non-living items do not possess life. However, it was a challenge for children to attribute life status to plants, which are inanimate living things.

Children’s persisting trouble is particularly noticeable when they are questioned about biological properties different from “alive” (e.g., Anggoro, 2006; Backscheider et al., 1993; Hatano, et al., 1993; Inagaki & Hatano, 1996; Opfer & Siegler, 2004; Springer & Keil, 1989;
Waxman, 2005). For example, Anggoro et al. (2008) demonstrated this difficulty by asking children to sort some cards representing both living and non-living items with respect to several predicates, such as “Is X alive?” “Can X grow?” and “Can X die?” Although they correctly classified living things when prompted with the predicates “grow” and “die,” they frequently disregarded plants when prompted with the word “alive.”

Leddon et al. (2008) believe that the word “alive” is unclear in English since it does not solely, or even chiefly, relate to the Western science-inspired biological explanation that research in this area emphasizes. Due to this lack of clarity, they suggest that even though children possess growing knowledge of an encompassing concept connecting human beings, non-human animals, and plants, utilizing the word “alive” conceals their comprehension of it. They further put forward that while this word especially applies to human beings, non-human animals, and plants, in practice, its utilization is commonly aligned to animate items only, thus excluding plants.

Since “alive” is commonly utilized in this animate sense, Leddon et al. (2008) believe that it may be linked to children’s hardship in considering plants as alive when questioned about this concept. They investigated 4- to 10-year-old children’s knowledge about living things in two experiments with two different probing terms: “alive” and “living thing.” In the first experiment, statistical analyses revealed that children attributed the word “alive” most often to animates, followed by plants and then non-living things. Children of all ages had no trouble linking life status to animate items and denying it to non-living things. However, they did not consider plants as things that are alive, implying that they misaligned the term “alive” with the concept “animate.” In the second experiment, the word “alive” was replaced with “living thing.” Statistical analyses revealed that children aged 4 to 7 tended to attribute the term “living thing”
most often to animates, followed by plants and then non-living things. Contrary to experiment one, 9- to 10-year-olds considered plants and animates as living things to the same extent. The authors deduced that children aligned the term “alive” to the concept “animate” rather than the concept “living thing.” Furthermore, substituting “alive” with “living thing” did not end the difficulty completely since it is only from age 6 to 7 that children reliably link life status to plants. They therefore suggested that this also shows why children are inclined to correctly assign life status to animate items and not to non-living things, but find it hard with inanimate living things such as plants.

Various studies performed at distinct age levels have demonstrated that the major reasons for participants’ interest in animals rather than plants are motion (Kinchin, 1999; Wandersee, 1986). The possible reasons for noticing animistic and anthropomorphic ideas in students are due to science teachers utilizing such kinds of explanations to describe certain topics (Watts & Bentley, 1991).

Subsequent research findings opposite to Piaget’s findings.

The meaningful description of children’s understanding of the concept of life has resulted in a considerable amount of research (Hatano et al., 1993). There are disagreements among developmental psychologists regarding young children’s abilities to build scientific explanations (Harris, 2010). Piaget (1930) believed that young children’s logical ability to build genuine scientific explanations was deficient, and that their capacity to produce scientific explanations developed throughout childhood into adulthood. Piaget declared that young children were inclined to give various non-scientific explanations for natural phenomena, such as animistic explanations. As Harris (2010) pointed out, post-Piagetian investigators state that Piaget misjudged young children’s abilities for scientific explanations.
Further disagreements regarding children’s ability to use one criterion or more for life surfaced. Piaget did not recognize the fact that children are able to utilize more than one criterion for life at one time (Laurendeau & Pinard, 1962). Piaget’s assumption that children’s comprehension of the life concept unfolded in stages was criticized by Carey (1985), who stated that Piaget’s complete data was not tabulated, but was rather in the form of chosen portions of interviews that represented the stages. Similar to Laurendeau and Pinard (1962), Carey found that, irrespective of age, children applied various sorts of justifications, for example, activity and growth, to explain their responses regarding why they considered certain things as living and others as non-living. Moreover, instead of relying only on clinical interviews, Venville (2004) studied children who were studying about living things in the classroom and the results reflected utilization of more than one criterion by a child.

Piaget’s interviews were repeated by Laurendeau and Pinard (1962) and Carey (1985). In both investigations, children initially were asked some orientating questions such as whether they knew the meaning of alive, after which they were shown various pictures and asked whether the items depicted in the pictures were alive or not and why. While Laurendeau and Pinard assessed a sample of 500 participants aged between 4 and 12, Carey assessed only 30 participants of three age groups, namely, 4-, 7- and 10-year-olds, and with 10 children in each age group. Although Carey’s sample was small, she indicated that her results largely reflected that of Laurendeau and Pinard. The results of both studies were represented in stages similar to Piaget’s results. It was found that children’s comprehension of the concepts of living and non-living things evolved gradually by 4 years of age, and that around the age of 10, more than half of the number of children believed that only animals and plants possessed life.
The attribution of animistic thought in children has been challenged in several instances (Bullock, 1985). Children show less animistic thought when questioned about objects to which they are accustomed (Berzonsky, 1971). Huang and Lee (1945) and Klingensmith (1953) observed that children 8- to 9-year-old frequently employed the term alive to inanimate items but did not accord these items with animistic properties to the same extent. The outcomes of these inquiries suggest that Piaget may have misjudged the extent of children’s animistic thought by counting on their knowledge of the meaning of the term alive and by assessing them about comparatively complicated or unfamiliar items (Bullock, 1985).

Studies considering preschool children’s comprehension of the way living things are distinct from artifacts has shown that children are able to differentiate between living things and artifacts. As Keil (1992, 1994) put forward, even 3-year-old children can distinguish that while the properties of biological things are related to an increase in survival, those of artifacts function to provide help to people. Novel research on this topic, done by Greif et al. (2006) and Kemler Nelson et al. (2001) has focused on a new way to study preschool children’s concept of living things. Their method emphasizes young children’s inquiries about artifacts and animals in contrast to prior studies that employed prompts to obtain children’s responses to particular questions. Exploring preschoolers’ questions about new items has shown that even 3-year-olds favor functional information when coming across novel artifacts and biological classification information when coming across novel animals (Greif et al., 2006).

Previous studies have shown that preschool children exhibit comprehension of the biological processes that are shared by living things, different from their comprehension of psychological processes, and they soundly differentiate between living things and artifacts in terms of the comprehension of the similarities among living things and the differences between
living things and artificial items (Backscheider et al., 1993; Erickson et al., 2010; Greif et al., 2006; Inagaki & Hatano, 1996; Jipson & Gelman, 2007; Keil, 1992, 1994; Kemler Nelson et al., 2001; Opfer & Gelman, 2001; Opfer & Siegler, 2004; Rosengren et al., 1991). In this regard, preschool children positively differentiate living from non-living things in terms of biological properties like death, growth, and regrowth, as well as in terms of the information they look for when coming across unfamiliar objects (Greif et al., 2006; Inagaki & Hatano, 1996; Waxman, 2005). For example, regarding growth and regrowth properties, young children rightly state that animals and plants grow (Inagaki & Hatano, 1996; Rosengren et al., 1991) and are able to recuperate by regrowth in case of damage (Backscheider et al., 1993). Regarding the death property, they rightly recognize that plants and animals die, but not artifacts (Nguyen & Gelman, 2002; Waxman, 2005). Furthermore, they correctly believe that plants and animals take in nutrients as a means of feeding themselves (Inagaki & Hatano, 1996, 2002).

Effect of culture and language on conceptions of living and non-living things.

Due to the assumption that culture may have an effect on biological understandings, research on such understandings have been performed in diversified societies, and it has been found that children’s notions in different societies do not always appear similar (Hatano et al., 1993). One example is the study conducted by Inagaki and Hatano (1987) with Japanese children, where most of the 5-year-olds declared that a tulip can feel pretty, indicating that they attributed to plants a property that is applicable to complex animals only. A considerable number of children also thought that the tulip can feel cold and/or happy. Another example is that of Stavy and Wax (1989), who found that 50% of their sample of Israeli 6- to 12-year-olds believed that plants were not alive, and a great number considered plants as neither living nor non-living things, but instead as an intermediate category.
It has also been found that children’s notions about the life status of inanimate items may also be different between societies and that this can be due to cultural and linguistic factors. An example that proposes that culture influences children’s reasoning about the concept of life is that preschool children in the United States scarcely ever assign life or other attributes that are unique to living things to any terrestrial inanimate item (Dolgin & Behrend, 1984; Gelman, Spelke, & Meck, 1983), while on the contrary, Japanese preschool children at times do attribute psychological properties to inanimate items with no motion or function, such as stones (Inagaki & Hatano, 1987).

It is worth examining the cultural and linguistic factors that cause Japanese children to show a greater tendency to consider plants and inanimate objects as alive and possessing attributes of living things. For example, Hatano et al. (1993) state that Japanese culture encompasses a belief that plants are to a great extent like humans. They also put forward that due to Japanese folk psychology, which includes possession of minds by inanimate items, a view of plants as having feelings and emotions, and a view of large and old inanimate items like mountains as being alive or even divine, Japanese children tend to display non-scientific ideas based on the behavior of their parents toward plants as well as large, ancient inanimate items.

Similar to cultural factors, linguistic factors can also contribute to Japanese children’s judgments. In this case, Hatano et al. (1993) refer to the use of the word “alive” and state that the colloquial expression corresponding to “alive” can be assigned to inanimate items only in a figurative sense; however, since the Chinese character corresponding to the concept kanji has typical meanings of “fresh” and “perishable” and also “alive,” therefore kanji is applied to substances like cake, wine, sauce and other inanimate but perishable substances. They further put forward that the same kind of attitudinal and linguistic factors may result in Israeli children being
less capable than U.S. or Japanese children of considering plants as being alive and having properties of living things. According to Stavy and Wax (1989), within Israeli traditions, plants are viewed as quite distinct from human beings and other animals in terms of possession of life. Furthermore, with respect to varying societal attitudes, while the Hebrew word for “animal” is quite close to that of “living” and “alive,” the word for “plant” has no apparent association to these words (Hatano et al., 1993).

In their study of Canadian and Israeli children from grades 2 to 6, Stavy and Wax (1992) found that Canadian children had significantly greater ability to classify plants as living things as compared to Israeli children. The authors suggest that the difficulty Israeli children face in categorizing plants as living things may be conceptual in origin. They further point out that language may influence concept development on verbal and perceptual levels. They believe that the use of the Hebrew language might have accounted for the results obtained because children may have been influenced by the logic inherent in their native language.

When putting the emphasis on language, some researchers have found that children’s attainment regarding the life status of items may be partly due to their inability to understand the word “alive” in a similar way to adults (Carey, 1985; Nguyen & Gelman, 2002; Piaget, 1929; Slaughter et al., 1999). According to Carey (1985), it is probable that these investigations were simply inquiring into children’s understanding of the word “alive” rather than their recognition of an encompassing concept connecting all living things. Due to their failure to comprehend that animals and plants are living things and that they form part of the same category, children’s persisting hardship is more commonly considered as proof that they find it hard to understand the meaning of the word alive (Slaughter et al., 1999).

**Problem**
Child assessment is an essential and expanding element of high quality early childhood programs. Epstein, Schweinhart, DeBruin-Parecki, and Robin (2004) view child assessment as valuable because it helps adults determine, comprehend, and support young children’s development and progress. They note that assessment can also be used to record and judge program effectiveness in meeting the educational needs of young children, determine staff development needs, and prepare future instruction.

There is concern regarding the way science is taught in schools in the United States (Duschl, Schweingruber, & Shouse, 2007) and considerable interest in the development of skills related to science during the preschool years (Duschl et al., 2007; Greenfield et al., 2009). Preschool curricula with foundational elements in science are being developed (e.g., French, 2004; Gelman, Brenneman, MacDonald, & Román 2010). However, examination of early childhood assessment in general (Snow & Van Hemel, 2009), and specifically science, (Brenneman et al., 2011) reveals a shortage of early science assessments that can be regarded as valid and reliable. Funding has been provided for the development of curriculum and assessment instruments in science at the early childhood level by the National Science Foundation and the U.S. Department of Education (French & Woodring, 2013).

Examples of instruments that are used to examine young children’s cognitive achievement include the Peabody Picture Vocabulary Test (PPVT), the Woodcock Johnson III (WJ-III) battery, and the Science Learning Assessment (SLA). A brief description of these instruments, including why they are not appropriate for the present study, follows.

Since there are insufficient developmentally appropriate tools to measure young children’s science learning, general standardized measures of cognitive progress like the PPVT have been employed to examine student achievement across teaching programs. The PPVT was
used in a study by French (2004) to assess the efficiency of a preschool science curriculum. While the PPVT and other similar measures provide an indication of vocabulary and cognitive abilities that may indirectly be promoted through teaching science, they do not provide enough information about the characteristic and magnitude of children’s science learning across programs (Samarapungavan, Mantzicopoulos, Patrick, & French, 2009).

The Science subtest of the Woodcock Johnson III (WJ-III) battery is a science-specific standardized measure of young children's science attainment (Woodcock, McGrew, & Mather, 2001). Three subtests, including social studies, humanities, and science, make up the Academic Knowledge cluster, with the science subtest focusing on the assessment of general knowledge in biological and physical sciences. Some of the science components are appropriate for preschool children. However, normative data are not specific to the Science subtest, but pertain to the whole Academic Knowledge cluster, and proof that the test is sensitive to teaching effects over time is unavailable (Samarapungavan et al., 2009). In addition, the test does not provide information about children’s domain-specific knowledge and competence across essential science themes and concepts (Schrank, McGrew, & Woodcock, 2001).

Another measure of science learning is the Science Learning Assessment (SLA), which assesses the science knowledge of kindergarten students. It consists of two subtests, namely, the Scientific Inquiry Processes subtest (not the focus of this present study) and the Life Sciences Concepts subtest, which evaluates understanding of concepts associated with living things and the physical environment. However, this assessment was specifically designed for students who are exposed to an inquiry-based science curriculum, which forms part of the Scientific Literacy Project (Samarapungavan et al., 2009). Furthermore, the SLA had been designed for kindergarten students, whereas the present study is focused on preschool children.
In addition to the reasons provided above, other early science interventions have employed assessments connected to specific interventions (e.g., Gropen, Clark-Chiarelli, & Hoisington, 2006; Samarapungavan, et al., 2007). This translates into the fact that there are insufficient pedagogically sensitive assessments that can be utilized to measure, gather, and examine young children’s learning outcomes, particularly across varying teaching methods (Samarapungavan et al., 2009). In order for assessments to be effective, they need to be practical, inexpensive, and meet rational norms of efficiency and validity (Epstein et al., 2004). Based on the review above, it is clear that measures of preschoolers’ competence in domain-specific science areas that are either directly or not directly connected to a particular curriculum are required. This is because a lack of developmentally appropriate and psychometrically strong instruments for measuring domain-specific science learning hinders research in this area.

To the researcher’s knowledge, there is an absence of reliable and valid assessment tools that can be used to measure, gather, and examine preschoolers’ competence related to basic life sciences, and that also provide normative data specific to the test rather than pertaining to a whole academic knowledge cluster. The assessment tool developed for this study has been designed in a way that can be employed either in the form of a formative assessment tool not directly connected to a particular curriculum or as a curriculum-based assessment tool. For the present study, the tool was considered as a curriculum-based assessment. The life sciences curriculum used at the child care centers where data was collected were covered in the form of one or more of the following: adaptation from Piaget’s work, shared book-reading, field trips, and nature observation and exploration.

Assessment in education: Curriculum-based measurement.
A fundamental but very challenging component of education is the measurement of change. According to Espin, Shin, and Busch (2005), educators are able to reliably assess student learning and the consequences of instructional interventions on that learning through the measurement of change in attainment. However, these scholars note that in spite of change measurement being essential, it is not the emphasis of educational assessment since the most common approach is the measurement of performance at only one point in time.

Regarding the availability of assessment tools, there is a measurement system that is termed curriculum-based measurement (CBM). CBM is regarded as a continuous data collection system aligned to classroom curriculum that is designed to provide teachers with information on students’ progress in thinking (Shinn & Bamonto, 1998) and on the consequences of instructional interventions on that progress (Espin et al., 2005) so that they can direct their planning and instructional decisions (Shinn & Bamonto, 1998). It is especially designed to measure change in student accomplishment by allowing for repeated measures within short time periods (Espin et al., 2005). Deno (1985) states that the measures established to be utilized as a component of CBM are simple, effective, easy to comprehend, and inexpensive, in addition to enabling recurrent measurement of student achievement over time.

The validity and reliability of CBM measures as indicators of achievement in the primary skill areas of reading, mathematics, spelling, and written expression for elementary school pupils have been supported by more than 25 years of research (Espin et al., 2005). Espin et al. (2005) reviewed information about correlations, test-retest reliabilities, alternate-form reliabilities, and change in student achievement. Regarding the correlations between CBM indicators and several criterion measures, the values typically ranged from .60 to .90 (Espin et al., 2005). As for test-retest and alternate-form reliabilities, the values were typically greater than .80 (Marston, 1989).
Regarding change in student achievement, several studies have demonstrated an improvement in student achievement when teachers have utilized CBM measures to assess and alter their instruction (Fuchs, Deno, & Mirkin, 1984; Fuchs, Fuchs, & Hamlett, 1989a, 1989b, 1989c; Fuchs, Fuchs, Hamlett, & Allinder, 1991; Fuchs, Fuchs, Hamlett, & Ferguson, 1992; Fuchs, Fuchs, Hamlett, & Stecker, 1990; Stecker & Fuchs, 2000; Wesson et al., 1988). Furthermore, CBM has also been integrated with statistical techniques, for example, hierarchical linear modeling, to produce student growth curves to be utilized to answer questions regarding the connection between student progress and instructional variables (Compton, 2000; Shin, Deno, & Espin, 2000). A study by Espin et al. (2005) confirmed that CBM can be utilized to identify the extent to which pupils’ performance is inconsistent to that of their peers at only one time point as well as to analyze the extent to which pupils are making progress relative to their peers, meaning that pupils who are inconsistent in performance as well as progress would be the ones who would require intensive interventions the most.

As pointed out by Moomaw, Carr, Boat, and Barnett (2010), it is challenging to assess young children in a reliable way because of the social and emotional elements associated with their developmental stage. Due to this, they further state that curriculum-based measures can be developed in a way that can be used more favorably with young children. The following table presents the difficulties associated with assessing young children as well as how the difficulties can be addressed with the use of curriculum-based measures.
**Table 1.**

**Challenges to Assessing Young Children, along with a Comparison of Traditional versus Curriculum-Based Assessments**

In contrast to traditional assessments, curriculum-based assessments align with classroom curriculum and offer teachers ongoing information on students’ development.

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Traditional Assessment</th>
<th>Curriculum-Based Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autonomy</td>
<td>Because preschool children are developing and asserting autonomy (Erikson, 1950), issues of compliance often arise during assessment procedures.</td>
<td>The measures can be developed to align with young children’s interests.</td>
</tr>
<tr>
<td>Attention</td>
<td>Preschool children typically have much shorter attention spans than school-aged children; this characteristic affects the length of time that can be devoted to assessing them.</td>
<td>Curriculum-based measures are designed to be much shorter than other types of developmental assessments.</td>
</tr>
<tr>
<td>Trust</td>
<td>Because preschool children are also establishing trusting relations (Erikson, 1950) outside their homes, fear about working with an assessor is a concern.</td>
<td>The measures can be administered by classroom teachers who have training on procedures and scoring.</td>
</tr>
<tr>
<td>Development &amp; Maturation</td>
<td>Maturation and learning occur quickly during the preschool years; therefore, ongoing assessment is critical for educational planning.</td>
<td>Curriculum-based measures are designed for making ongoing instructional decisions.</td>
</tr>
<tr>
<td>Alignment to Curriculum</td>
<td>Assessment information is often not related to standard preschool experiences and curricula.</td>
<td>Assessment is directly aligned to curriculum.</td>
</tr>
</tbody>
</table>

Significance

Current research plans (e.g., Anderson & Helms, 2001; Rennie, Feher, Dierking, & Falk, 2003) require action for addressing science literacy in the early school years for two reasons: first, to prepare children sufficiently for learning science when they reach upper grades, and second, to maintain a continuing interest in learning about science (Mantzicopoulos et al., 2008). Seen in this light, it is considered worthwhile to identify (Ekici et al., 2007), explore (Tunnicliffe et al., 2008), and evaluate (Morrison & Lederman, 2002) the ideas that students bring to class in order to facilitate learning of new material. Since prior knowledge is the foundation upon which new knowledge is built (Meyer, 1993), and since misconceptions are to a certain degree resistant to change (Ekici et al., 2007; Özay & Öztas, 2003) and may thus persist in adulthood, studying preschool children’s scientific thinking and reasoning is a worthwhile step in order to address misconceptions before more complex parts of a topic are taught. Students can thus be helped to fulfill a common goal of science education, which is developing meaningful understanding of scientific concepts and knowing how to use these concepts in everyday life (Kara & Yeşilyurt, 2008), that is, becoming scientifically literate (Türkmen & Usta, 2007).

Studying pupils’ understanding of basic concepts such as the concept of life and living things is a key step in determining how biology teaching is being performed (Yorek et al., 2009) and how it should be performed. Since children usually express their mental models through conversations, drawings, and writing, teachers should take advantage of this and assess students' knowledge and understanding (Tunnicliffe, 2000) so as to be better informed of any misconceptions they may have. Harris (2010) stated that science teachers are especially interested in examining the role of student-generated explanations in science learning since various researchers have proposed that explanation has an important role in learning (Chi,
Bassok, Lewis, Reimann & Glaser, 1989; Wellman & Lagattuta, 2004). In spite of the disparities that exist in the literature regarding children’s abilities to build scientific explanations, explanation is a frequent occurrence, and learning to give explanations, including scientific ones, is a valuable part of young children’s linguistic and conceptual development in its own right (Harris, 2010). Furthermore, since many studies show that children have typical beliefs about natural phenomena that may inform practice (Driver, Squires, Rushworth, & Wood-Robinson, 1994), it is fundamental to find out what children think and their reasoning as such if science teaching is to boost children’s science learning (Osborne & Freyberg, 1985). Therefore, the instrument used in this study examined preschool children’s understanding of and ideas about basic life sciences through conversations in the form of questions from assessors and answers from children.

The significance and requirement of investigating children’s thoughts and comprehension as a means for advising curriculum development and instructional design have been put forward by Driver and colleagues. According to Driver, Guesne, and Tiberghien (1985a), children’s ideas affect their learning in various ways, such as the observations they make, the sense they make from these observations, and the methods they employ to build new ideas and understandings. Driver et al. also note that through comprehension of children’s ideas, curriculum development and instructional design can be better adjusted to children due to information obtained regarding the science concepts to be taught, the learning experiences that would challenge children’s notions and understandings, and the aims the activities would assist.

The content-specific ideas and understandings children have regarding particular phenomena and events are necessary to curriculum development and instructional design (Driver, Guesne, & Tiberghien, 1985b). As a result, these scholars conclude that curriculum
development and instructional design must take into account the conceptual structure of the subject as well as the students’ notions. In addition, comprehending children’s ideas leads to the development of assessments that better determine and examine children’s conceptual understanding (Shepardson, 2002). In line with this study, identifying children’s ideas about basic life sciences may help in developing curriculum and instructional activities as children’s developing understandings would be taken into account. The significance of this study lies in the development of an assessment to contribute to the compilation of knowledge regarding children’s ideas about the life sciences. Understanding children’s ideas about basic life sciences will enrich the knowledge base about children’s biological thinking. By understanding children’s ideas about the life sciences, possible hindrances to learning can be determined and insight obtained toward designing curriculum and instruction that builds on children’s understandings. Based on this, conduction of the study may provide valuable information about the designing of tools to measure children’s conceptions of life sciences. Seen in this light, developing and validating an assessment tool focusing on measuring preschoolers’ conceptions of basic life sciences in terms of the concept “living thing” can be considered as a significant contribution for early childhood science education. Also, it is hoped that after validation of the instrument, teachers will be able to use it and be informed of children’s conceptions of basic life sciences so that they can plan instruction accordingly for young children. Therefore, this study is only for validation purposes with the intent of future studies based on the outcome of this validation study.

The findings of this study will be helpful for teachers, textbook authors, and designers of science curricula as well as preschool children themselves. As described below, teaching can be
improved when teachers use a validated instrument to know the misconceptions that students hold or are likely to hold (Kara & Yeşilyurt, 2008).

- Because misconceptions are resistant to change, if preschool children’s misconceptions are known, teachers can help them make necessary changes to their mental opinions in order to reach appropriate conclusions (Modell et al., 2005). In other words, teachers can help students recognize the errors that arise due to misconceptions and encourage them to either reject or alter them (Wescott & Cunningham, 2005).

- Teachers may get a better idea of the amount of time to spend on a particular topic, of the appropriate amount of details to go into when teaching a particular topic (Meir, Perry, Herron, & Kingsolver, 2007), and of the areas to emphasize when teaching a particular topic (Bulunuz et al., 2008).

- Teachers may have a better idea of how to stop teaching in ways that reinforce misconceptions (Dikmenli, 2010; Meir et al., 2007).

- Helping preschool children in their learning about basic life sciences enables them to proceed to the learning of more complex and abstract concepts of that topic more easily when they move to higher grades.

- Teachers as well as textbook authors may be more careful about the terminology they use during instruction and book writing respectively (Dikmenli, 2010) in order to avoid reinforcing misconceptions.

- If students’ difficulties in understanding certain concepts are taken into consideration, science curricula can be developed in a way that is more likely to decrease or avoid misconceptions and lead to better understanding and meaningful learning of a topic (Özay & Öztas, 2003).
To summarize, it can be said that the development of adequate learning experiences for children requires a curriculum that takes into consideration children’s ideas and the expected science learning aims (scientific perspective), enabling curriculum and instruction to proceed in such a way that it leads children toward scientific understanding and curricular continuity (Driver et al., 1994). This can be achieved, in part, by using CBMs directed at measuring change either in science achievement in general or in specific topics.

**Hypothesis**

The working hypothesis pertaining to this study is as follows: The four sub-constructs, each of which can be measured through a series of questions on the LSA, will make a significant contribution to the latent construct of scientific thinking and reasoning, thereby causing the LSA to be a significant indicator of preschool children’s conceptions of basic life sciences. The four sub-constructs are “Naming”; “Classifying (classification: animal, plant, or something different, CL_apsd)”; “Living (classification: living or non-living, CL_lnl)”; and “Explaining (scientific explanation: living or non-living, SE_lnl).” In the context of this study, scientific thinking and reasoning is defined as the act of using one’s mind to understand and draw particular logical judgments and conclusions about living things that include animals and plants, and non-living things, from a scientifically acceptable perspective.

**Research Questions**

This study attempted to answer two research questions as mentioned below.

Research Question 1: What is the contribution of each sub-construct to the latent construct on the LSA?

Research Question 2: What is the technical adequacy of the LSA in measuring preschool children’s conceptions of basic life sciences?
Chapter 2: Methods

The purpose of this study was to design and validate the Life Sciences Assessment (LSA). Data collection for this study took place over six weeks, during November and December of the 2012-13 academic year.

Sampling

The target population for this study was all preschool children aged 3 to 5 years enrolled in licensed preschools or child care programs in Hamilton County in the state of Ohio. The accessible population for this study was all preschool children aged 3 to 5 from the greater Cincinnati metropolitan area. The convenience sample considered for the study was all children enrolled at four local child care centers. Parents were informed of the purpose and objectives of the study by means of a letter that had been approved by the University of Cincinnati Institutional Review Board. Only children whose parents gave consent for their participation were selected to participate. The data collection for this study spanned six weeks. Out of a total of 250 parent permission slips distributed for the main study, 146 signed permission slips were returned to the classroom teachers. Of these 146 children, only 124 completed the assessment. The other 22 children did not take part in the study either due to absenteeism, involvement in other activities during the period of data collection, for example, field trips, or simply refusal to participate. Thus, the final sample size for this study was 124 children.

 Participant inclusion criteria.

Participants’ inclusion criteria were based on the following: first, aged 3 to 5 years; second, English as primary spoken language; and third, a signed consent form from a parent/legal guardian allowing him/her to take part. Even if a child’s parent/legal guardian provided permission for the child to participate in the study, the child was not forced to
participate if he/she did not wish to do so. The demographic breakdown of the participants for the main study is provided table 2 below.

Table 2

Demographic Information of Participants for Main Study

<table>
<thead>
<tr>
<th>Demographic information</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>67</td>
</tr>
<tr>
<td>Females</td>
<td>57</td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>3 years</td>
<td>56</td>
</tr>
<tr>
<td>4 years</td>
<td>65</td>
</tr>
<tr>
<td>5 years</td>
<td>3</td>
</tr>
<tr>
<td>Ethnicity/Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>81</td>
</tr>
<tr>
<td>African-American</td>
<td>29</td>
</tr>
<tr>
<td>Asian</td>
<td>10</td>
</tr>
<tr>
<td>Hispanic</td>
<td>4</td>
</tr>
<tr>
<td>Socio-economic level</td>
<td></td>
</tr>
<tr>
<td>Lower-income families</td>
<td>42</td>
</tr>
<tr>
<td>Middle to upper-income families</td>
<td>82</td>
</tr>
<tr>
<td>English as second language</td>
<td>0</td>
</tr>
<tr>
<td>Physical and/or cognitive impairment</td>
<td>0</td>
</tr>
</tbody>
</table>

Instrumentation

The measurement tool for this study was a curriculum-based assessment termed the Life Sciences Assessment (LSA), designed to measure young children’s conceptions of basic life sciences (Appendix A). It consisted of a series of photographs with four questions related to each
photograph. This format was selected since researchers tend to interview children after showing them pictures, photographs, or flashcards (e.g. Barman et al., 2000; Barman et al., 2006; Gatt et al., 2007; Tunnicliffe et al., 2008). Colored photographs were used in this study to provide participants with visual cues.

**Instrument design.**

The instrument was piloted twice with five children in each test. It was revised after each pilot test before being used in the main study. Details about the instrument, the procedures undertaken for the main study, as well as the pilot testing and revision, are provided in the following sections.

**Photographs included in the LSA.**

The LSA included nine photographs (Appendix B): three animals – fish (B1), butterfly (B2), and dog (B3); three plants – strawberry (B4), grass (B5), and flower (B6); and three non-living things – snowman (B7), airplane (B8), and clouds (B9). Most photographs were taken by the researcher and a family member. However, the photographs of the airplane (V. Kratochvil, n.d.) and the strawberry (P. Kratochvil, n.d.) were obtained from a website that allows their use for any purpose including commercial. The items included in the instrument were chosen to be those that are familiar to the children since they are commonly found in their immediate environment. With respect to their categories, the items were chosen based on their differences in physical appearances. The animals included a mammal, a fish, and an insect. The plants included a fruit, a flower, and green and well-trimmed grass. The non-living things included two items that show movement (a scientifically acceptable characteristic of living things) but are not living (airplane and clouds); and another item that shows characteristics of living things in cartoons but which is actually not living (snowman).
The majority of the items used in the instrument had previously been used in other studies:

- dog (Tunnicliffe et al., 2008)
- fish (Anggoro et al., 2008; Babai et al., 2010; Barman et al., 2000; Chen & Ku, 1997; Chen & Ku, 1998; Wax & Stavy, 1987)
- butterfly (Barman et al., 2000; Tamir, 1997; Villalbi & Lucas, 1991)
- flower (Babai et al., 2010; Barman et al., 2006; Chen & Ku, 1997; Gatt et al., 2007; Stavy & Wax, 1992; Wax & Stavy, 1987)
- grass (Barman et al., 2006; Chen & Ku, 1997; Gatt et al., 2007; Kwon, 2003; Stavy & Wax, 1992; Venville, 2004; Wax & Stavy, 1987)
- clouds (Anggoro et al., 2008; Chen & Ku, 1997; Chen & Ku, 1998; Salleh et al., 2007; Stavy & Wax, 1992; Tamir, 1997; Venville, 2004)
- airplane (Kwon, 2003)

Each photograph was printed in color on a sheet of paper that measured 8.3 x 11.7 inches. Each photograph was bordered by the white color of the sheet of paper. They were individually kept in page protectors and were randomized by means of an online tool (Haahr, 1998) before being administered to the children. They were presented to all children in the same order: snowman, fish, strawberry, clouds, butterfly, flower, grass, dog, and airplane.

**Instrument questions.**

Four questions targeting the children’s scientific thinking and reasoning were asked for each of the nine photographs. The specificity of the fourth question depended on the child’s answer to the third question. The questions are listed in order below.
1) “What is this?”

2) “Is a/an (correct name of item shown) an animal, a plant, or something different?”

3) “Is a/an (correct name of item shown) a living thing?”

4) “What makes you think a/an (correct name of item shown) is a living/non-living thing?”

**Goals of the LSA.**

The LSA comprised of four main goals that included measuring:

1. preschool children’s ability to identify a photographed item by correctly stating its common name
2. preschool children’s ability to distinguish the photographed item as an animal, a plant, or something different
3. preschool children’s ability to distinguish whether the photographed item was living or non-living
4. preschool children’s scientific explanation regarding why the photographed item was considered as living or non-living

The items measuring these goals are organized in a Table of Specifications (Appendix C).

**Measurement scale.**

In terms of scoring, the LSA follows a type of summative response scale. In other words, each response related to a specific question on the LSA can be analyzed independently, or can be summed with other related items in order to end with a score for a group of statements. The numbers on each scale are arranged in an ascending order, thus reflecting higher performance on the concept being measured. While a scoring range of 0 to 2 has been attributed to the first three questions, a range of 0 to 3 has been attributed to the last question. The same range was applicable to each photograph. Regarding the first three questions, a score of 2 was attributed for
a completely correct response; a score of 1 was attributed for a correct answer, but where a child showed uncertainty; and a score of 0 was attributed in cases of no response or an incorrect answer. Regarding the last question, 1 point was given for each scientifically acceptable criterion, with a maximum of 3 points if three or more acceptable answers were provided. No points were given in cases of no response or incorrect answers. The scientifically acceptable characteristics of living things expected as answers to this question on the LSA include the seven criteria put forward by King and Sullivan (as cited in Venville, 2004, p. 458). The criteria include reproduction, growth, nutrition, use of energy, movement, sensitivity, and excretion. A point to note in the scoring of this last question was about the uncertainty in deducing whether a child correctly mentioned scientifically acceptable criteria, for example, a child saying that the dog is a living thing because it has eyes did not indicate that the child understood that dogs possess senses. Venville (2004), who came across similar responses while studying kindergarten children’s ideas about living things, indicated that a score of 0 was to be attributed to such ambiguous cases. Thus, such cases were scored as 0 in the present study. Details regarding the scoring guide and the assessment guide are provided in Appendices A2-A6.

*Pilot testing and revision.*

The LSA was tested in two pilot studies with five different children in each pilot test. Both tests were performed in September of the 2012-13 academic year. The researcher was the only assessor who collected data from the 10 children, implying that the same skills were used in interacting with and assessing every child. Some changes were made to the instrument after each test. The first pilot test, which included 12 photographs and five related questions, resulted in the following changes: 1) removal of three photographs, namely, bird for the animal category, trees for the plant category, and rocks for the non-living things category, in order to decrease the
length of the measure due to signs of distraction and fatigue in children; 2) changing the wordings (“something else” to “something different,” and “How can you tell…” to “What makes you think…”) of some questions (refer to table 3 below) to create less confusion in children; and 3) discarding one question (refer to table 3 below) due to the confusion created by the words “How can you tell…” and “something else,” in addition to the fact that it seemed difficult for the children to provide an explanation for an answer chosen from three options in the previous question.

Following these revisions, the LSA was again pilot tested with five children. The only change consisted of replacing the word “it” by the correct name of the item shown, in the second, third, and fourth questions, in order to ensure that a child was focusing on the current photograph shown rather than still thinking about the previous photograph(s). Even though the two pilot tests were conducted with 10 children only, the information obtained was sufficient to make a decision about the length as well as the wording for the final instrument.

The LSA questions for the main study were based on the feedback obtained from the two pilot tests as well as information obtained from the literature. The questions were phrased in a way that reflected what had been performed in previous studies. Examples of studies which employed the question “What is this?” were carried out by Venville (2004) and Leddon et al. (2008). Examples of studies which employed the question “Is (item name) a living thing?” were carried out by Wax and Stavy (1987), Stavy and Wax (1992), and Chen and Ku (1997). Examples of studies which employed the question “Is it an animal?” or “Is the (item name) an animal?” were carried out by Venville (2004) and Villalbi and Lucas (1991) respectively. However, there is a slight addition in wording in the current study since the goal is for the children to identify whether the photograph shown represents an animal, a plant, or something
different. The question “What makes you think that a/an (item name) is/is not a living thing?” was phrased based on previous studies by Wax and Stavy (1987), Stavy and Wax (1992), and Villalbi and Lucas (1991), who started their question with “What makes you think…?” In case a child responded “I don’t know” to this question for any photograph, the following prompt, which is based on a previous study, was used: “Can you tell me anything about it?” (Bullock, 1985).

The table below shows the evolution of the questions from the first pilot test to those used in the main study.

Table 3

*Evolution of Questions from First Pilot Test to Main Study*

<table>
<thead>
<tr>
<th>First pilot test</th>
<th>Second pilot test</th>
<th>Main study</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is this?</td>
<td>What is this?</td>
<td>What is this?</td>
</tr>
<tr>
<td>Is it an animal, a plant, or</td>
<td>Is it an animal, a plant, or</td>
<td>Is a/an (correct name of item shown)</td>
</tr>
<tr>
<td>something else?</td>
<td>something different?</td>
<td>an animal, a plant, or something different?</td>
</tr>
<tr>
<td>How can you tell it’s an animal/a</td>
<td>Discarded</td>
<td>Discarded</td>
</tr>
<tr>
<td>plant/something else?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is it a living thing?</td>
<td>Is it a living thing?</td>
<td>Is a/an (correct name of item shown) a living thing?</td>
</tr>
<tr>
<td>How can you tell it’s a living/non-</td>
<td>What makes you think</td>
<td>What makes you think a/an (correct name of item shown) is a living/non-</td>
</tr>
<tr>
<td>living thing?</td>
<td>it’s a living/non-living thing?</td>
<td>name of item shown) is a living/non-living thing?</td>
</tr>
</tbody>
</table>
Instrument administration.

Two assessors were involved in the data collection process for the main study. Training that included explanations and discussions of procedures involved in the study was provided by the principal investigator to the second assessor.

Procedure prior to instrument administration.

The assessors worked from a list of children whose parents/legal guardians had returned signed informed consent forms. Each assessor made contact with the child to be assessed within his/her classroom, established rapport, and then walked with the child into an assigned room close to the classroom. Although preschool children may not communicate verbally, they do give clear indications (through facial expressions, body language, gestures, and so forth) when they do not want to participate in an activity. Therefore, verbal communication as well as the indications mentioned above was taken into consideration during interaction with the children.

The assessor helped the child get seated and then asked him/her to play a science game, “Hi (child’s first name). My name is (assessor’s name). This is a game about science where you get to name some objects and talk about them. OK?” When the child agreed (orally, or through facial expressions, or body language such as nodding head, or other gestures, and so forth), the assessor said, “Thank you, (child’s first name). Let us start with the first picture.”

Procedure during instrument administration.

All children were shown the photographs and asked the questions in the same pre-arranged order. Each child was assessed only once individually during the study period. Administration of the instrument, as well as the scoring criteria, was standardized. At the conclusion of instrument administration, the assessor thanked the child and walked him/her back to the classroom.
Data Management

Data file structure.

There were three sources of data in this study: 1) demographic information; 2) qualitative data from the participants’ comments, and 3) quantitative data from the scoring of participants’ responses. Demographic information, that is, age, gender, socioeconomic status, and race/ethnicity of children, which were obtained from the school’s records upon receipt of parental consent, are the only identifiers which were made accessible to the principal investigator. The demographic information was entered into an Excel file. Assessment data, which included the participant’s study identification number (used instead of the child’s name on the study questionnaire), the date of assessment, the assessor’s initials, the child care center at which the participant was enrolled, the time taken to complete the assessment, qualitative data from the participants’ comments, and the participant’s scores on each item, were entered into a separate Excel file.

Data quality.

Data coding.

Four types of data required coding: 1) participant name; 2) gender; 3) ethnicity; 4) and socioeconomic level. Each name was given a separate identification number starting from 001 to 124. Gender was coded as 0 for male participants and 1 for female participants. Ethnicity, as obtained from the school’s files, was self-reported by families. According to the National Institutes of Health (2001) guidelines, the ethnicities reported fell into four categories, namely White, Black, Asian, and Hispanic, and were given a numeric code of 1 to 4 respectively. Socioeconomic level was coded as 0 for the children from middle to upper-income families and 1 for the children from lower-income (as determined by Head Start) families.
Data checking.

Assessment data that was entered into an Excel file was also entered into a Statistical Package for the Social Sciences (SPSS) 21.0 file. In order to ensure accuracy of this data, two checks were performed. First, the principal investigator reviewed each data point on the scoring sheets and checked them manually after input into the SPSS file. Any data entry errors were corrected. Second, SPSS frequencies were computed to confirm that there were no values out of range.

Data security.

Information about participants, including children’s identities and research data, has been and will be kept confidential. Only the principal investigator has access to the research data. Any research data that was saved on a computer and transferred to a flash drive as back-up has been password-protected. Should the data be published, no child will be identified by name, classroom, teacher, or school at any time.

Data and assumption checks.

Data checking procedures were implemented using SPSS 21.0 to verify the quality of the data and to check univariate and multivariate assumptions. The dataset was evaluated for the following: missing data, normality, linearity, and outliers.

Missing data.

Missing demographic or assessment data were not applicable to the study. All demographic information required was obtained from the child care centers. No assessment data was missing because instrument administration and scoring were performed by the assessors themselves. However, missing data arose at the individual student level with respect to the individual item response, that is, when a child either did not respond to a specific item or
indicated that he/she did not know the answer. In such cases, a score of zero was attributed to corresponding items.

**Normality.**

The assumption of normality is met when a frequency distribution follows a normal distribution. A normal distribution is symmetrical, with the highest frequency found in the middle, and frequencies decreasing when moving toward each extreme (Gravetter & Wallnau, 2009). Skewness and kurtosis values for a normal distribution are zero (Tabachnick & Fidell, 2007). According to Waigandt (2004), it is a common belief that skewness and kurtosis values within the limits of $\pm 1.0$ give a suitable indicator of normality. Normality was assessed both statistically (skewness and kurtosis values, and Kolmogorov-Smirnov test) and visually (histogram) by using SPSS 21.0 with the main study data. A point to note is that the distributions for all four scales of the LSA are censored at both ends. This implies that a child cannot score below zero on any question, or above two on any question for the first three scales, or above three on any question for the last scale. Therefore scores do not taper off at the ends of the distributions.

**Skewness and kurtosis.**

Descriptive statistics, including mean, standard deviation, skewness, and kurtosis, were calculated (Appendix D, Table D1). Skewness and kurtosis values were within acceptable limits of $\pm 1.0$ for two of the four scales, namely, “Classifying” and “Living.” While the skewness and kurtosis values for the scale “Classifying” were -.362 and -.980 respectively, the skewness and kurtosis values for the scale “Living” were -.068 and .379 respectively. In regards to the scale “Explaining,” the kurtosis value was still within the acceptable limit at .581. However, the skewness value of this scale was 1.236, that is, slightly above the limit. As far as the scale
“Naming” is concerned, the skewness value was -1.787 and the kurtosis value was 3.045, indicating a violation of normality. Finally, the skewness and kurtosis values for the total score on the LSA were -.208 and -.881 respectively. These values, which were within acceptable limits of ±1.0, clearly indicate a normal distribution.

*Kolmogorov-Smirnov (K-S) test.*

Kolmogorov-Smirnov (K-S) values were also calculated for each of the four scales, as well as for the total measure of the LSA (Appendix D, Table D2). K-S values for each individual scale were significant, indicating a possible violation of normality: “Naming” ($p = .000$), “Classifying” ($p = .002$), “Living” ($p = .003$), and “Explaining” ($p = .000$). However, the K-S value for the total measure was not significant ($p = .111$), indicating a normal distribution.

*Histograms.*

Visual assessment of normality in the form of histograms (Appendix E, Figure E1) by using SPSS 21.0 with the main study data was performed. The histogram for the scale “Naming” showed the lowest peaks at the left hand side of the distribution (scores of 12, 13) and the highest peaks close to the mean (17.15) at the right hand side of the distribution (scores of 16, 17, 18), indicating that a large portion of the sample (113 children) were able to correctly provide the common name of the items shown to them. The highest peak represented a score of 18, which was the maximum score that could be obtained for this scale. This indicated that almost two thirds (77 children) of the sample were able to correctly provide the common name of all photographs shown. High scores, which were expected for the majority of the children, can be explained by the fact that photographs shown were familiar to them. Asking for the common name of the photographs implied probing into socially transmitted knowledge and therefore was
not found as too challenging. On the whole, the pileup of scores to the right of the distribution and the long left tail indicated a negatively skewed distribution.

The histogram for the scale “Classifying” showed that about two thirds of the sample obtained scores that were centered around the mean of the distribution. The highest peak, representing a score of 16, was found to the right hand side of the histogram, indicating that many children obtained scores above the mean. The column for the highest score (score of 18) indicated that few children (6.5%) were able to correctly classify all items shown to them into animal, plant, or something different. A low column completely to the left hand side of the distribution represented a score of 2, indicating that there were no children who were not able to classify any item into animal, plant, or something different. On the whole, with about two thirds of the scores located in the middle of the distribution and a slight negative skewness (-.362), the distribution was close to normal.

The histogram for the scale “Living” showed that the highest peaks (scores of 10 and 12) were found in the middle of the distribution and very close to the mean (10.67). Over two thirds of the sample obtained scores that were found in the middle of the distribution. A very low frequency at the complete right hand side of the histogram (score of 18) showed that very few children (3.2%) were able to correctly classify all photographs shown into living or non-living things. A very low frequency at the complete left side of the histogram (score of 3, 4, and 5) showed that some children found it challenging to correctly classify photographs shown into living or non-living things. On the whole, over two thirds of scores were located in the middle of the distribution and there was very low negative skewness (-.068), indicating a distribution that was very close to a normal distribution.
The histogram for the scale “Explaining” showed that as scores increased, frequency decreased. The highest frequency (score of 0) was found completely at the left hand side of the distribution and indicated that two-fifths of the sample did not yet demonstrate scientific reasoning abilities regarding why photographs shown were living or non-living. The lowest frequency (score of 7) was found completely at the right hand side of the distribution, indicating that very few children (2.4%) were able to provide some scientifically acceptable characteristics regarding why photographs shown were living or non-living. However, even though a score of 7 was the highest score obtained for this sample, it was still a low score as compared to the maximum score of 27 that could be obtained for this scale on the LSA. The low scores obtained on this scale, hence producing a slightly positively skewed distribution (+1.236), were expected due to the young age of this sample. As reported in the background section, previous studies indicated that many children associate life with motion due to animistic thinking during early childhood. They also ascribe life to things that show some kind of activity or usefulness. Furthermore, Piaget (1930) believed that young children’s logical ability to build genuine scientific explanations was deficient, and that their capacity to produce scientific explanations developed throughout childhood into adulthood.

The histogram for the total score on the LSA showed that the highest peaks were found around the middle of the distribution and indicated that nearly two thirds of the sample obtained scores that were located in the middle of the distribution. The lowest frequencies occurred at the two ends of the distribution. The lowest frequency on the left end indicated that no child was completely unable to correctly address any questions on the LSA. The lowest frequency on the right end indicated that only few children scored rather high on the LSA. On the whole, nearly
two thirds of the sample obtained scores in the middle of the distribution, with a slight negative skewness (-.208), indicating a distribution that was very close to normal.

**Linearity.**

The presence of a straight-line relationship between two variables (where one or both variables can be combinations of various variables) indicates the assumption of linearity (Tabachnick & Fidell, 2007). The assumption of linearity was assessed both statistically and visually by examining Pearson correlations and scatterplots respectively. Using the main study data, Pearson correlations were examined to describe the degree of correlation, (a) among the scales on the LSA (Appendix D, Table D3), and (b) between age (in months) and the total score of each of the four scales on the LSA as well as the total score for the measure (Appendix D, Table D4). $R^2$ values (Appendix D, Table D4) were also examined to describe the percent of variance explained by the model. Scatterplots (Appendix E, Figure E2) comparing the children’s age in months to the total score of each of the four scales on the LSA as well as the total score for the measure, were also examined.

**Correlations among scales.**

Pearson correlations showed weak to strong linear relationships between pairs of scales. Correlation coefficients, $r$, ranged from .195 to .535. Weak correlations were obtained between the scales “Naming” and “Living” ($r=.195$) and “Naming” and “Explaining” ($r=.285$). Moderate correlations were obtained between the scales “Classifying” and “Living” ($r=.422$), “Classifying” and “Explaining” ($r=.449$), and “Living” and “Explaining” ($r=.476$). Strong correlations were obtained between the scales “Naming” and “Classifying” ($r=.535$). These data indicated that the assumption of linearity was met for all scales.

**Correlation between age in months and scores on the LSA.**
Correlation between age and the scores for two scales, “Classifying” and Explaining,” were low at \( r = .297, p = .001 \) and \( r = .281, p = .003 \) respectively, and were both statistically significant at the .01 level (2-tailed). Both scales approached a moderate correlation of \( r = .3 \). This denoted that an increase in age was significantly related to an increase in scores on these scales. This also showed that 8.8% and 7.9% of these two respective scales was explained by age. Correlation between age and the scores for the other two scales, “Naming” \( (r = .122, p = .176) \) and “Living” \( (r = .166, p = .065) \) were low and not statistically significant, indicating that these two scales were not explained by age. Finally, the correlation between age and the total score on the LSA was significant with a moderate value of \( r = .306 \), indicating that 9.3% of the total score was explained by age.

In addition to determining the degree to which age (in months) was correlated with each of the scales on the LSA as well as with the total measure by means of the Pearson correlation coefficient \( r \), scatterplots for each of these scales and the total measure were also generated (Appendix E, Figure E2). Compared to coefficient \( r \), scatterplots were helpful in determining the strength and direction of the relationship between these variables visually.

The scatterplot for the scale “Naming” showed a weak positive linear trend with some scatter. High scores (16 to 18) were obtained irrespective of children’s age. This was expected and can be explained by the fact that photographs shown were familiar to the children. Asking for the common name of the photographs implied probing into socially transmitted knowledge and therefore was not found as too challenging. However, age for this sample was not a predictor of naming ability, implying that there is not much scope to further develop young children’s ability to name familiar objects.
The scatterplot for the scale “Classifying” showed a weak positive linear trend with a good deal of scatter. Low to high scores were obtained irrespective of age in months. This finding can relate back to the literature in that children’s understanding of the concept animal is more restricted than that used in science (Tunnicliffe et al., 2008). Likewise, children hold a limited meaning of the word plant as compared to the one used in science (e.g. Bell, 1981a; Gatt et al., 2007).

The scatterplot for the scale “Living” showed a weak positive linear trend with a good deal of scatter. Low to high scores were obtained irrespective of age in months. However, age for this sample was not a predictor of preschool children’s ability to classify photographs shown into living or non-living things. As reported in the background section, previous studies indicate that many children associate life with motion due to animistic thinking during early childhood. Piaget (1930, 1951) found that children up to six years of age ascribed life to things that show some kind of activity or usefulness.

The scatterplot for the scale “Explaining” showed a weak positive linear trend with some scatter. Regardless of age in months, all children scored rather low for this scale. Heavily stacked scores at the value of zero implied that regardless of age, many children were not able to provide any scientifically acceptable answers. This finding can relate to Piaget’s (1930) belief that young children’s logical ability to build genuine scientific explanations is deficient, and that their capacity to produce scientific explanations develop throughout childhood into adulthood.

Finally, for the total score on the LSA, the scatterplot showed a moderate positive linear trend with a good deal of scatter. Low to high total scores were obtained irrespective of age.

Outliers.
An outlier is a score or measurement that illustrates an atypical exception to a general pattern by diverging largely from other scores or measurements in a group (Fraenkel & Wallen, 2009). It can be a univariate outlier, that is, an extreme value on one variable, or a multivariate outlier, that is, a strange combination of scores on more than one variable (Tabachnick & Fidell, 2007).

**Univariate outliers.**

A check for univariate outliers was performed by calculating the values for each of the four scales and for the total scores for the LSA, using the formula $M \pm 4 \times SD$ since the sample size was greater than 80 (Pan, 2011). All values fell within the calculated range, indicating absence of univariate outliers (Appendix D, Table D5).

**Multivariate outliers.**

A check for multivariate outliers was performed by calculating Mahalanobis’ Distance, $\chi^2_{.001}(4)_{\text{critical}} = 18.47$, for the four scales on the LSA. No values fell above the critical $\chi^2 (18.47)$, indicating absence of multivariate outliers.

**Data Analysis Plan**

The data analysis plan included addressing the two research questions in the following way:

1) **Research Question 1:** “What is the contribution of each sub-construct to the latent construct on the LSA?” was addressed by confirmatory factor analysis (CFA), using Analysis of Moment Structures (AMOS) 21.0 (Arbuckle, 2012).

2) **Research Question 2:** “What is the technical adequacy of the LSA in measuring preschool children’s conceptions of basic life sciences?” was addressed by reliability assessments (internal consistency reliability using SPSS 21.0 and construct reliability)
and validity assessments (content validity, construct validity (CFA), and internal validity).

**Analysis of first research question.**

CFA involves specifying and figuring one or more satisfactory models of factor structure where each model puts forward a series of latent variables (also known as factors) to explain covariances among a group of observed variables (Doll, Raghunathan, Lim, & Gupta, 1995). A CFA using AMOS 21.0 was performed to address the first research question. Specifically, it was performed to measure the level of contribution of each sub-construct to the latent construct on the instrument, as well as the model fit.

Creating hypothesized models was the first step in data analysis. These models were based on logic, theory, and previous studies. After running a CFA with these models, one final best model was considered for this study. This hypothesized model (Appendix E, Figure E3) consisted of four sub-constructs as measured by the four scales of the instrument. Therefore, contributions of the four sub-constructs of this model to the latent construct were assessed as well as the model fit. In other words, the CFA was used, first, to examine the relationships between the scales on the instrument and the latent construct, and second, to determine model fit by examining fit indices.

Contributions of sub-constructs to the latent construct were established by examining factor loadings, also known as beta weights (Kline, 2000) or regression coefficients (Hox & Bechger, 2007) or regression weights. Factor loadings, which represent the correlations of the variable with the factor (Kline, 2000), were examined for statistical significance at $\alpha = .001$ as well as practical significance. Regarding model fit, Tabachnick and Fidell (2007) note that this is an active area of research. They believe that the comparative fit index (Bentler, 1990) and the
root-mean-square error of approximation (Browne & Cudeck, 1992) are the most commonly reported fit indices. In addition to these two fit indices, Thompson (2004) states that two other fit indices, namely the chi-square ($\chi^2$) statistical significance test and the normed-fit index (Bentler & Bonnett, 1980) are also among the most frequently used today to assess the goodness of fit between the hypothesized and measured models. A rough criterion in determining model fit, as suggested by Tabachnick and Fidell, is the ratio of the $\chi^2$ to the degrees of freedom. Thus, these five fit indices were taken into consideration for this study. Their definitions are provided below.

- $\chi^2$ statistical significance test: test of model's potential to reproduce the sample variance/covariance matrix (Doll et al., 1995)
- Normed-Fit index (NFI): comparison between the $\chi^2$ value for the estimated model against the $\chi^2$ value for the independence model while assuming that the measured variables are totally independent (Thompson, 2004)
- Comparative Fit index (CFI): assessment of model fit relative to a baseline null or independence model (Thompson, 2004)
- Root-Mean-Square Error of Approximation (RMSEA): determination of how well model parameters will do at reproducing population covariances (Thompson, 2004) or more simply, assessment of the lack of fit in a model as compared to an excellent (saturated) model (Tabachnick & Fidell, 2007)
- Ratio of $\chi^2$ to degrees of freedom: designates the level of efficiency of competing models in accounting for the data (Doll et al., 1995).

**Analysis of second research question.**

Reliability assessments and validity assessments were taken into account to address the second research question: What is the technical adequacy of the LSA in measuring preschool
children’s conceptions of basic life sciences? The reliability of the LSA was assessed by considering internal consistency reliability and construct reliability. In terms of validity, content validity, construct validity, and internal validity were considered. Content validity was addressed in stages. Construct validity was established by the CFA. Potential threats to internal validity – history, maturation, and location, were controlled. However, test-retest reliability, inter-rater reliability, and criterion validity were not performed for reasons discussed below.

Reliability assessments.

Reliability designates the consistency of scores or answers from the administration of one instrument to another, and from one group of items to another (Fraenkel & Wallen, 2009). In other words, a reliable instrument is one that provides consistent results. Training that included explanations and discussions of procedures involved in the study was provided by the principal investigator to the second data collector in order to maximize the standardization of the data collection process. The reliability of the LSA was assessed through internal consistency reliability and construct reliability.

Internal consistency reliability.

Internal consistency refers to the strength of the relation of the items within a scale to the latent variable (DeVellis, 2003). In other words, it is the consistency of each participant’s scores from item to item. Coefficient alpha (Cronbach, 1951), also known as Cronbach’s alpha, is viewed as the best indicator of reliability in terms of internal consistency (Kline, 2000). In addition, White (2011) observes that the most frequently reported measurement of internal consistency for quantitative measurement tools is Cronbach’s alpha. The internal consistency of the LSA was assessed through calculation of Cronbach’s alpha for all the items of the LSA using SPSS 21.0.
Construct reliability.

The second form of reliability considered was construct reliability (Fornell & Larcker, 1981). According to Hamdan, Badrullah, and Shahid (2011), construct reliability is calculated to determine the consistency of construct validity indicators. In other words, construct reliability indicates the degree to which the indicators are internally consistent. The formula below (Hamdan et al., 2011) was used to calculate the construct reliability of the LSA.

\[
\text{Construct reliability} = \frac{\text{square of total standardized loadings}}{\text{square of total standardized loadings} + \text{measurement errors}}
\]

where measurement error = 1 - (standardized loading)^2

Reliability not performed.

Two reliability assessments that were not performed include test-retest reliability and inter-rater reliability. The initial plan included assessment of reliability through test-retest reliability. The test-retest method involves administering the same instrument a second time to the same individuals after a certain interval of time has elapsed (Fraenkel & Wallen, 2009). To demonstrate trustworthiness, there should be at least a 3-month gap between administration of the two tests, as well as samples consisting of at least 100 participants (Kline, 2000). However, the reliability coefficient is likely to be lower with a longer time interval due to a greater possibility of changes in subjects, which thus involves the need to choose an appropriate time interval (Fraenkel & Wallen, 2009). Conducting a test-retest for this study, while simultaneously satisfying the conditions of minimum number of participants as well as appropriate time interval, proved to be difficult. First, closing of one child care center for the summer period resulted in loss of nearly 50% of participants. Second, due to the busy summer period for the other child care centers, the child care center staff did not get an opportunity to redistribute parent
permission slips for the retest. Since seven months had already elapsed, it was decided that the
test-retest would not be carried out.

In addition to not having done the test-retest reliability, inter-rater reliability also was not
calculated. In the case of the LSA, the first three sets of questions were straight-forward as well
as the responses obtained. Hence, no training was needed. However, the last set of questions
required expert knowledge in order to be scored properly. Therefore, instead of computing inter-
rater reliability for 25% of participants, a biology content expert with extensive classroom
experience was contacted to verify the categories obtained for the last set of questions and the
number of points attributed.

Validity assessments.

Validity refers to whether the deductions made by the investigator regarding the data
obtained by using an instrument are appropriate, meaningful, correct, and useful (Fraenkel &
Wallen, 2009). In other words, Fraenkel and Wallen (2009) refer to the validity of an instrument
as the degree to which results from it allow researchers to make guaranteed inferences regarding
the characteristics of the individuals under study. In terms of the validity of the LSA, content
validity, construct validity, and internal validity were examined. However, due to reasons
discussed below, it was not possible to examine criterion validity. Information regarding the
three types of validity as they pertain to this study is provided below.

Content validity.

Content validity designates judgments on the content and reasonable structure of an
instrument as it is to be utilized in a particular study (Fraenkel & Wallen, 2009). Scale
development of the LSA was addressed in stages before being presented to professionals to
account for content validity. First, an attempt to define the latent construct theoretically was
made. Since the life sciences encompasses content that can be overwhelming to include in a single assessment tool designed for preschool children, it was decided to narrow down the assessment to basic life sciences that only include scientific thinking and reasoning regarding particular criteria of living things, such as animals and plants, and non-living things that are naturally-occurring and man-made. A sub-construct that theoretically does not form part of the life sciences construct but that was deemed important to include in this assessment tool involved probing into children’s socially transmitted knowledge that would lead to correct identification of particular photographed items. This sub-construct was included to ensure that children were aware of what they were looking at because they had to respond to questions about these photographs. Following the decision of what would be measured by the latent construct, possible photographs reflecting the three categories (animal, plant, and artifact) intended to be included in the LSA were brainstormed and noted. Next, possible questions regarding knowledge of basic life sciences were brainstormed and noted. Following this, the literature was searched to determine which photographs were previously used in other studies and to thus narrow the selection down to 12 photographs (4 in each category). The literature was also referred to in order to word the five questions in a way appropriate for preschool children. Photographs to be included were then either taken with a digital camera or obtained through a web search. Following this initial process, the instrument was pilot tested, revised, and pilot tested a second time. The updated instrument (after the second pilot test) and a definition of what was to be measured were presented to four professionals for feedback. These professionals possess one or more of the following characteristics: experience in teaching science at preschool level, experience in developing science curriculum for preschool level, knowledge in science literacy
matters related to early childhood education, and experience with statistical and assessment measures.

*Construct validity.*

In the field of psychology and education, the term construct is used to refer to unobservable traits that are determined by the researcher to sum up or to explain regularities or relations in observed behavior (Thorndike & Thorndike-Christ, 2010). Construct validity designates the nature of the psychological construct or attribute being measured by the instrument (Fraenkel & Wallen, 2009). In other words, Fraenkel and Wallen (2009) refer to construct validity as the extent to which the entire evidence obtained is consistent with theoretical anticipation. A CFA was performed to measure the level of contribution of each sub-construct to the latent construct on the LSA. In other words, the CFA enabled the examination of the relationships between the scales on the instrument and the latent construct.

*Internal validity.*

A study is said to have internal validity when any relationship detected between two or more variables is directly related to the intended variables instead of being the result of something else (Fraenkel & Wallen, 2009). One or more possible hypotheses are available to explain the outcomes of a study that lacks internal validity. Researchers refer to these possible hypotheses as threats to internal validity (Fraenkel & Wallen, 2009). The main potential threats to internal validity that could have hindered the process of this study were assumed to be history, maturation, and location. Therefore, ways to minimize them were conceived. These potential threats as well as the ways devised to minimize them are described below.

The history threat occurs when on occasion, one or more unexpected and unprepared situations may happen while a study is in progress that can affect the responses of subjects
(Fraenkel & Wallen, 2009). History was identified as a potential threat because during the course of the study, children may be exposed to and taught about items or similar items (such as models or actual specimens) to be used in the study, as part of their regular curriculum. In order to address the history threat, teachers were not informed about what was to be assessed in the study. They were asked to allow children to play with any authorized animals, plants, or non-living things (models or actual specimens) that they wished to during the course of the study, but were asked not to interact with the children from a life sciences perspective during the process.

The maturation threat occurs when alterations during the study may be due more to factors related to the passing of time rather than to the study itself (Fraenkel & Wallen, 2009). In the case of maturation, very young children mature rapidly, and any change in their knowledge about living things, such as animals and plants and non-living things, may be due to aging and experience. The maturation threat was minimized because the study took place during school hours and spanned a short period of time, that is, 6 weeks.

The location threat occurs when an assessment is performed in particular locations that may cause alternate reasons for results (Fraenkel & Wallen, 2009). Location was considered as a potential threat because different locations for data collection may create different responses in children. Since data was collected from four different child care centers, the location threat was addressed within the child care centers. Although there were four different data collection locations, there was only one in each child care center. These locations involved few distractions and were the same for all participants of a particular day care center.

*Criterion validity.*

Criterion validity designates the relationship between the scores obtained from one instrument and the scores obtained with one or more similar instruments or measures (Fraenkel
& Wallen, 2009). Fraenkel and Wallen (2009) refer to this type of validity as the extent to which information obtained with one instrument is in agreement with information obtained from the use of other independent instruments. Criterion validity was not examined in this study because of the lack of appropriate assessment measures for this young age group (Thorndike & Thorndike-Christ, 2010).
Chapter 3: Results

Research Question 1: What is the contribution of each sub-construct to the latent construct on the LSA?

A CFA was performed to answer this research question. In order to assess the contribution of the four scales on the LSA, the hypothesized model (Appendix E, Figure E3) was evaluated through AMOS 21.0. Based on the modification indices suggested in the AMOS output, one covariance was added to the model between the scales “Living” and “Explaining” (Appendix E, Figure E4).

All goodness of fit indices that were examined indicated a good fit for the model. The $\chi^2$ test was not significant, $\chi^2 (1, N = 124) = 0.929$, $p = .335$. Values for NFI, CFI, and RMSEA were .992, 1.000, and .000 respectively. The ratio of the $\chi^2$ to the degrees of freedom was 0.929. All standardized regression coefficients (Appendix E, Figure E4) demonstrated statistical significance at $p < .001$ (2-tailed). Since they ranged from .43 to .98 ($\beta > .3$), they also demonstrated practical significance. The standardized regression coefficients and their corresponding $R^2$ values were as follows: “Naming,” $\beta = .55$, $R^2 = .30$; “Classifying,” $\beta = .98$, $R^2 = .96$; “Living,” $\beta = .43$, $R^2 = .18$; and “Explaining,” $\beta = .46$, $R^2 = .21$, where $R^2$ is the percent of variance in the indicator variable that is explained by the factor (Tabachnick & Fidell, 2007), or more specifically, as pertaining to this study, the percent of variance in the sub-constructs that is explained by the latent construct.
Research Question 2: What is the technical adequacy of the LSA in measuring preschool children’s conceptions of basic life sciences?

Technical adequacy is a common term that provides a description of the levels of reliability and validity of a particular assessment (Hampton, 2011). The different levels of reliability examined for the LSA included internal consistency reliability and construct reliability. The different levels of validity examined for the LSA included content validity, construct validity, and internal validity.

Reliability.

Reliability was assessed in terms of internal consistency reliability and construct reliability.

Internal consistency reliability.

The internal consistency reliability of the LSA was assessed through calculation of Cronbach’s Alpha for all the items of the LSA using SPSS 21.0. The resulting Cronbach’s Alpha was .733.

Construct reliability.

The construct reliability was calculated by using the formula below. The resulting value was .714.

\[
\text{Construct reliability} = \frac{\text{square of total standardized loadings}}{\text{square of total standardized loadings} + \text{measurement errors}}
\]

where measurement error = 1 - (standardized loading)^2

Validity.

Validity was assessed in terms of content validity, construct validity, and internal validity.

Content validity.
Content validity was addressed in stages, as described in the data analysis plan: theoretical definition of latent construct; brainstorming about photographs and related basic life sciences questions for preschool children; examining related literature; obtaining photographs; first pilot testing; revision of the assessment tool; second pilot testing; and finally presentation of the assessment tool to professionals for feedback. Since there were no suggestions for change from the committee of experts, the instrument was considered as valid in terms of content validity.

**Construct validity.**

A CFA was performed to measure the level of contribution of each sub-construct to the latent construct on the LSA. In other words, a CFA was used to examine the relationships between the scales on the instrument and the latent construct scientific thinking and reasoning. The factor loadings for each scale were as follows: “Naming,” $\beta = .55$; “Classifying,” $\beta = .98$; “Living,” $\beta = .43$; and “Explaining,” $\beta = .46$. All standardized regression coefficients (Appendix E, Figure E4) demonstrated statistical significance at $p < .001$ level (2-tailed). Since they were all greater than .3, they also demonstrated practical significance.

**Internal validity.**

As previously indicated, the main potential threats to internal validity that could have hindered the process of this study were assumed to be history, maturation, and location. These three potential threats to internal validity were within the control of the researcher to some extent. Ways to minimize them during the course of the study were successful.
Chapter 4: Discussion

The purpose of this study was to design and validate an assessment tool termed the Life Sciences Assessment (LSA) in order to measure preschool children’s conceptions of basic life science concepts. Specifically, it was hypothesized that four sub-constructs that make up the scales of the instrument, “Naming,” “Classifying,” “Living,” and “Explaining,” would make a significant contribution to the latent construct of scientific thinking and reasoning, thereby causing the LSA to be a significant indicator of preschool children’s conceptions of basic life sciences. This study examined the contribution of each of these four scales as well as the technical adequacy of the LSA.

Major Findings

The hypothesis that the four sub-constructs making up the scales of the LSA are significant contributors to the latent construct of scientific thinking and reasoning and thus significant indicators of preschool children’s conceptions of basic life sciences was supported by this research. The CFA indicated a good fit between the hypothesized and measured models. Moderate to high factor loadings were obtained. The factor loadings demonstrated statistical as well as practical significance. The technical adequacy indicated that the LSA is a reliable and valid measure for this sample of preschool children.

First research question.

CFA was used to evaluate the model fit between the hypothesized and measured models. The initial evaluation of the hypothesized model suggested the addition of a covariance between the scales “Living” and “Explaining.” The model was significantly improved with the addition of this covariance. This improvement was represented by the following changes, which provided an indication of a good-fitting model.
1) change in the $\chi^2$ test from significant to non-significant: A non-significant $\chi^2$ is desired, thus indicating no significant difference between the hypothesized and observed relationships (Tabachnick & Fidell, 2007).

2) increase in the value of NFI from .881 to .992: Regarding the NFI, values of .95 or greater represent reasonable model fit (Thompson, 2004).

3) increase in the value of CFI from .893 to 1.000: Regarding the CFI, values of around .95 represent reasonable model fit, with values close to 1.000 desirable (Thompson, 2004). In this study, the LSA achieved the best desired value.

4) decrease in the value of RMSEA from .213 to .000: An RMSEA value of .000 shows that a model was estimated to accurately reproduce the population covariances. In the case of the RMSEA, reasonable model fit is represented by values of around .06 or less (Thompson, 2004). In this study, the model was estimated to accurately reproduce the population covariances.

5) decrease in the ratio of the $\chi^2$ to the degrees of freedom from 6.597 to .929: A ratio of the $\chi^2$ to the degrees of freedom of less than 2 is an indication of a good-fitting model (Tabachnick & Fidell, 2007).

The next step was examination of factor loadings for statistical and practical significance. Examination of the AMOS output showed that all factor loadings demonstrated statistical significance at $\alpha = .001$ (2-tailed). Regarding the practical significance, since factor loadings higher than .3 can be considered as significant (Kline, 2000) and that a factor loading of over .3 is the minimum level of practical significance (Hair, Black, Babin, Anderson & Tatham, 2006), all factor loadings in the measured model demonstrated practical significance. The higher the loading, the more the variable is a pure measure of the factor (Tabachnick & Fidell, 2007).
indicated that the scales represented the latent construct in the following increasing order: “Living” (β = .43), “Explaining” (β = .46), “Naming” (β = .55), and finally “Classifying” (β = .98). Thus, the latent construct explained low to high percentages of variance in the subconstructs in the following increasing order: “Living” (R² = .18), “Explaining” (R² = .21), “Naming” (R² = .30), and finally “Classifying” (R² = .96).

Second research question.

The technical adequacy, as examined by internal consistency, construct reliability, content validity, construct validity, and internal validity, indicated that the LSA is a reliable and valid measure for this sample of preschool children, as indicated below.

Internal consistency reliability.

Regarding the internal consistency reliability, according to DeVellis (2003), research scales with coefficient alpha below .60 is unacceptable; between .60 and .65 is undesirable; between .65 and .70 is minimally acceptable; between .70 and .80 is respectable; between .80 and .90 is very good; and greater than .90 requires a decrease in the length of the scale. Since coefficient alpha for the LSA was .733, the internal consistency reliability was regarded as respectable.

Construct reliability.

Regarding construct reliability, the desired value should be .7 or higher (Hair et al., 2006). With a value of .714, the LSA demonstrated construct reliability.

Content validity.

As for content validity, which was addressed in stages, since there were no suggestions for any change from the committee of experts, the instrument was considered as valid in terms of content validity.
Construct validity.

Construct validity, as established by the CFA, resulted in factor loadings ranging from .43 to .98, thus demonstrating practical significance since all factor loadings were higher than .3 (Hair et al., 2006) as well as statistical significance at $\alpha = .001$ (2-tailed).

Internal validity.

The three potential threats to internal validity, namely, history, maturation, and location, were within the control of the researcher to some extent. Since ways to minimize these potential threats during the course of the study were successful, the LSA is considered as internally valid.

As discussed above, it is clear that the hypothesis that the four sub-constructs making up the scales of the LSA are significant contributors to the latent construct of scientific thinking and reasoning, and thus significant indicators of preschool children’s conceptions of basic life sciences, was supported by this research. This was indicated by the following:

1) good model fit as demonstrated by the most commonly reported fit indices (non-significant $\chi^2$; NFI = .992; CFI = 1.000, RMSEA = .000; ratio of $\chi^2$ to degrees of freedom = .929);

2) respectable internal consistency (Cronbach’s alpha = .733)

3) desired value for construct reliability (.714)

4) content validity addressed in stages, as required

5) construct validity as established by CFA – high factor loadings that demonstrate statistical significance of factor loadings at $\alpha = .001$ (2-tailed) as well as practical significance of factor loadings (all factor loadings being greater than .3)

6) successful minimization of potential threats to internal validity
Thus, the LSA can be considered as a reliable and valid measure of the conceptions of basic life sciences of this sample of preschool children.

**Other findings**

In regards to naming the photographs, it was expected that the children would be able to name them easily since they were familiar with the items and the question simply probed into socially transmitted knowledge. Over 93% of children were able to correctly provide the common name of all the photographs except the snowman. The percentage of accurate naming ranged from 93.5% to 98.4%, implying that on the whole, there was no photograph that was correctly identified by all the children. A surprising finding regarding the snowman was that only 96 children, that is, 77.4% of the sample, could readily identify it. Although the snowman was dressed up to look like one, 15 children, that is, 12.1% of the sample, simply identified it as snow.

Regarding the classification of the photographs into animal, plant, or something different, the most challenging animal to classify was the fish, which was correctly classified by about 68% of the sample only. The butterfly was correctly classified by about 77% and the dog by 75% of the sample. The most challenging plant to classify was grass, which was correctly classified by about 35% of the sample. The strawberry was correctly classified by about 47% and the flower by about 72% of the sample. The most challenging non-living thing to classify was the snowman, which was correctly classified by about 65% of the sample. The clouds were correctly classified by about 80% and the airplane by about 76% of the sample. In the case of the animals, not reaching 100% of correct classification can be explained by the fact that children’s understanding of the concept animal is more restricted than that used in science (Tunnicliffe et al., 2008); for example, irrespective of age, students usually utilize the word animal in a way that is analogous to the biological concept of mammal (e.g. Bell, 1981b). In this study, some children
may have considered the fish as a comparable set rather than a subset of animals, as was previously found (e.g. Salleh et al., 2007; Villalbi & Lucas, 1991). The butterfly also had previously been classified as non-animal (Bell, 1981b; Villalbi & Lucas, 1991). It is possible that some children in this sample could also be viewing animals as large and terrestrial (Tunnicliffe et al., 2008) as well as possessing fur (Barman et al., 2000), three characteristics that do not reflect those of a fish or a butterfly. In the case of the plants, not reaching 100% of correct classification can be explained by the fact that children hold a limited meaning of the word plant as compared to the one used in science and that their mental model of plants reflects mostly the anatomic features and cultural outlook. Previous studies have indicated that students do not regard an organism as a plant unless it possesses particular features or parts, such as a flower or stem (Bell, 1981a), or stalk and leaves (Gatt et al., 2007). This finding may explain why the flower, whose leaves and stem were clearly visible, reached the highest percentage of correct classification among the plants, followed by the strawberry, whose leaves and stem were not as conspicuous. As for grass, children may have considered it as a separate group from plants rather than a sub-group of plants, as had previously been the case in Chen and Ku’s (1997) study. Aside from the absence of responses from children regarding how to classify the snowman, clouds, and airplane, some children identified these photographs as animals or plants, with animals selected to a greater extent than plants. This misconception can be related to the fact that the snowman shows life-related activities similar to those of animals in cartoons; that the shape of clouds change every day, or more than once in a day, and can look like animals; and that the airplane flies like a bird, which is actually an animal.

In the case of the classification of the photographs into living or non-living things, the degree of challenge in classifying the three animals as living things was almost equal, with about
78% of children correctly classifying them. As for the plants, the degree of challenge increased from the flower (about 56%) to the strawberry (about 48%) and finally to the grass (about 42%), thus causing the grass to be the most challenging to classify as a living thing. Regarding the non-living things, the percentage range of correct classification was small. The degree of challenge increased in the following order: clouds (about 51%), airplane (about 50%), and snowman (about 47%), thus causing the snowman to be the most challenging non-living thing to be correctly classified. The incorrect classification of the photographs by many children can be explained by the fact that life is associated with motion due to animistic thinking during early childhood, an idea that was brought up by Piaget (1929). Wrongly classifying the clouds and airplane can be explained by the idea that life is associated with motion. This idea can also apply to the snowman, which shows movement and other life-related activities in cartoons. This idea can also explain why the photographed animals that show a great degree of movement, reached the highest percentage of correct classification. Piaget found that children up to 6 years of age ascribed life to things that show some kind of activity or usefulness, and thus did not consider trees and other plants as living things. Several other studies have revealed that students tend to classify plants as non-living things (e.g. Anggoro et al., 2008; Hatano et al., 1993; Margett & Witherington, 2011; Yorek et al., 2009). In Yorek et al’s (2009) study, the students tended to mention animals as living things to a greater extent than plants, a finding applicable to this study as well.

The question that proved the most difficult for children was the one requiring reasons as to why the photographs represented either living or non-living things. Piaget (1930) believed that young children’s logical ability to build genuine scientific explanations was deficient, and that their capacity to produce scientific explanations developed throughout childhood into adulthood.
Findings revealed that three scientifically acceptable criteria were attributed to only one photograph, the fish, and by one child only. Only one scientifically acceptable criterion was obtained for three photographs, namely, the snowman, clouds, and the airplane, and this by a limited percentage of the sample (6.5% for the snowman and clouds and 2.4% for the airplane). An interesting point to note is that these three photographs depicted non-living things. Some children were able to provide at least two scientifically acceptable criteria for all the living things. Another interesting finding is that children were able to ascribe scientifically acceptable criteria to the animals to a larger extent than to the plants. In general, the scientifically acceptable characteristics mentioned by the children pertained to the categories of movement, nutrition, reproduction, growth, excretion, and sensitivity.

The questions probing into the children’s scientific knowledge reveal a greater number of correct answers attributed to animals than to plants. This finding leads back to Wandersee and Schussler’s (1999) term “plant blindness,” that is, the failure to see or notice plants in the environment. Furthermore, Patrick and Tunnicliffe (2011) have noted that teachers are responsible for causing students to hold misconceptions about plants when they spend more instructional time on the animal kingdom than on the plant kingdom. Together with the animistic thinking that prevails during early childhood and plant blindness, the amount of instructional time could have been a cause for the results obtained in this study.

**Limitations**

There are certain limitations that apply to this study. First, test-retest reliability was not performed. Although the initial plan included assessment of reliability through test-retest, it was not possible to do so in an appropriate time frame with an appropriate number of participants. Second, although the total LSA meets normality requirements, as determined both visually and
statistically, one individual scale, “Explaining,” shows a skewness value slightly above the limit, and another scale, “Naming,” shows a skewness value below the accepted limit and a kurtosis value above the accepted limit. However, a non-normal distribution was expected for the scale “Naming” since this scale was simply probing into socially transmitted knowledge, which was not challenging. Third, the sample included nearly double the number of children from middle-to upper-income families as compared to those from lower-income families, representing a ratio of approximately 2:1. This ratio indicates that this sample under-represents children from families with income above the poverty line and over-represents children from families with income below the poverty line; the percentage of families whose income was below poverty level in Hamilton county for the year 2012 was about 16% (Hamilton County, Ohio Demographics, 2013).

Implications

A number of pedagogical implications linked to scientific misconceptions arose from the use of the LSA in this study. In science classes, irrespective of grade level, teachers are faced with the major responsibility of addressing students’ alternative conceptions, which have been shown to come from the majority of science areas (Babai et al., 2010). A way for teachers to either avoid or address scientific misconceptions is by staying informed about children’s previous knowledge and the misconceptions they hold or are likely to hold. This implies that teachers should determine what and how students think about a particular topic, as well as find out about any conceptual or reasoning difficulties experienced by students. Furthermore, by staying informed about scientific misconceptions, teachers can assist students in recognizing the errors that arise due to misconceptions held and encourage students to either reject or alter them. Children’s responses to the LSA assessment can also be a way for teachers to be aware of
whether they are teaching in ways that are reinforcing scientific misconceptions instead of addressing them. In addition, knowledge of children’s scientific misconceptions can encourage teachers to emphasize areas more prone to be held as misconceptions by students when planning lessons. Also, teachers as well as textbook authors, can pay more attention to the terminology they use during instruction and book writing respectively. Finally, if students’ difficulties in understanding certain concepts are taken into consideration, science curricula can be developed in a way that is more likely to decrease or avoid misconceptions and lead to better understanding and meaningful learning of a topic. Use of the LSA can be a starting point to help teachers with their responsibility of addressing students’ alternative conceptions, such as classifying animals as non-animals (e.g. fish), classifying plants as non-plants (e.g. grass), classifying non-living items as animals (e.g. clouds) or plants (snowman), considering animals shown as a comparable set rather than a subset of animals (e.g. fish), considering plants shown as a comparable set rather than a subset of plants (e.g. grass), classifying living things as non-living things (e.g. strawberry) and vice versa (e.g. airplane), and providing non-scientifically acceptable criteria to describe why items are living or not (e.g. flower is living because it is pink, or snowman is not living because it has a hat). Based on the types and amount of misconceptions obtained, early childhood educators can plan instruction accordingly in order to devote sufficient instructional time to basic life sciences.

**Considerations for future research**

A number of considerations for future research emerged from this study. First, since test-retest reliability could not be performed, future use of the LSA could include retesting at least 100 participants after a 3-month interval to determine if there is a significant correlation between the initial testing and the retest. Second, parallel versions of the LSA could be developed and
tested. As such, repeated probes would be available for ongoing measurement. Third, a shorter version of the LSA that includes only two of each category of photographs, that is, two animals, two plants, and two non-living things, could be explored. This would lower the assessment time for busy teachers and may also cause less fatigue for children. Fourth, since there is a huge number of living and non-living things on earth, the LSA could be redesigned to include photographs of items different from those used in this study. It can be helpful to choose items based on what has previously been used in the literature. Fifth, the LSA could be redesigned to include fewer photographs but more life sciences related questions for each photograph in order to probe deeper into children’s scientific thinking and reasoning. Sixth, instead of asking questions verbally, the LSA could be made more interactive by asking children to classify the photographs into animal, plant, or something different, and then into either living or non-living thing. Seventh, in order to find out if there is a significant difference depending on the way the items are shown, the items could be videotaped and shown to children on a computer screen instead of using photographs. Finally, the LSA could be used with children attending different child care programs in different geographical areas. The children’s performance on the assessment could be helpful in determining the effectiveness of the programs. Teachers using the LSA could also determine if the tool informs their curricular activities.

Conclusion

The purpose of this study was to design and validate the Life Sciences Assessment (LSA) in order to assess preschool children’s conceptions of basic life sciences. The hypothesis that the four sub-constructs, each of which can be measured through a series of questions on the LSA, will make a significant contribution to the latent construct of scientific thinking and reasoning, thereby causing the LSA to be a significant indicator of preschool children’s conceptions of life
sciences, was confirmed. CFA, which was used to address the first research question, revealed that all the goodness of fit indices that were examined indicated a good fit for the model, and factor loadings were moderate to high. In addition, all standardized regression coefficients demonstrated statistical significance at $p < .001$ (2-tailed) as well as practical significance. The reliability assessments, namely internal consistency reliability and construct reliability, and the validity assessments, namely content validity, construct validity, and internal validity, that were used to examine the technical adequacy of the instrument indicated that the LSA is a reliable and valid measure for this sample of preschool children. Other findings revealed that the questions probing into the children’s scientific knowledge represent a greater number of correct answers attributed to animals as compared to plants, implying that in order to decrease the level of plant blindness, teachers should not favor a particular kingdom during instruction, but should instead spend sufficient amounts of instructional time on the different kingdoms. Furthermore, teachers can use the LSA to help them with their responsibility of addressing students’ alternative conceptions about animals, plants, living things, and non-living things. With a strong background about the basic life sciences that is free from scientific misconceptions dating from preschool, children can take the challenges of learning more advanced life sciences concepts in higher grades.
References


Conezio, K., & French, L. (2002). Science in the preschool classroom: Capitalizing on children’s fascination with the everyday world to foster language and literacy development. *Young Children, 57*, 1-5.


Wesson, C., Deno, S., Mirkin, P., Maruyama, G., Skiba, R., King, R., & Sevcik, B. (1988). A causal analysis of the relationships among on-going measurement and evaluation, the


### Appendix A

**Life Sciences Assessment**

Table A1

<table>
<thead>
<tr>
<th><strong>Student ID:</strong></th>
<th><strong>Center:</strong></th>
<th><strong>Date:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Assessor:</strong></th>
<th><strong>Classroom:</strong></th>
<th><strong>Time taken:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Directions</th>
<th>Target Answer and Scoring</th>
<th>Participant’s Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Engagement:</strong> Approach the child and say that you would like to play a game with him/her.</td>
<td>Not applicable</td>
<td></td>
</tr>
<tr>
<td><strong>Game orientation:</strong> Sit across from the child with the instrument including questions and photographs in a file. Assessor says, “Hi (child’s first name). My name is (assessor’s name). This is a game about science where you get to name some objects and talk about them.”</td>
<td>Not applicable</td>
<td></td>
</tr>
<tr>
<td>1. Identification #1: <em>(photograph #1: snowman)</em> <em>(a)</em> Assessor shows the photograph of a snowman and asks, “What is this?”</td>
<td><strong>SCORE:</strong> 0, 1, 2&lt;br&gt;Target answer = snowman&lt;br&gt;&lt;br&gt;(I) “snowman” (score = 2)&lt;br&gt;(II) “snow/ice” (no score, continue with prompt – “What has been made with the snow/ice?”)&lt;br&gt;(III) Any partially correct answer, for example, snow (after prompt), ice (after prompt), etc (score = 1)&lt;br&gt;  + REFER TO SCORING GUIDE #7</td>
<td></td>
</tr>
<tr>
<td><em>(b)</em> Assessor asks, “Is a snowman an animal, a plant, or something different?”</td>
<td><strong>SCORE:</strong> 0, 1, 2&lt;br&gt;Target answer = something different&lt;br&gt;&lt;br&gt;(I) “something different” (score = 2)&lt;br&gt;“animal/plant” (score = 0)&lt;br&gt;  + REFER TO SCORING GUIDE #1 and #7</td>
<td></td>
</tr>
</tbody>
</table>
| (c) Assessor asks, “Is a snowman a living thing?” | SCORE: 0, 1, 2  
Target answer = no  
(I) “no” (score = 2)  
“yes/sometimes” (score = 0)  
+ REFER TO SCORING GUIDE #4 and #7 |
| --- | --- |
| (d) Assessor asks, “What makes you think that a snowman is/is not a living thing?” | SCORE: 0, 1, 2, 3  
Target answer = 3 scientifically acceptable answers  
+ REFER TO SCORING GUIDE #6 and #7 |
| 2. Identification #2: (photograph #2: fish)  
(a) Assessor shows the photograph of a fish and asks, “What is this?” | SCORE: 0, 1, 2  
Target answer = fish  
(I) “fish/goldfish” (score = 2)  
(II) “animal” (no score, continue with prompt – “How is this animal called?”)  
(III) Any partially correct answer, for example, shark, etc (score = 1)  
+ REFER TO SCORING GUIDE #7 |
| (b) Assessor asks, “Is a fish an animal, a plant, or something different?” | SCORE: 0, 1, 2  
Target answer = animal  
(I) “animal” (score = 2)  
“plant/ something different” (score = 0)  
+ REFER TO SCORING GUIDE #2 and #7  
**Note:** Still provide this question if child said “animal” in question 2(a) |
| (c) Assessor asks, “Is a fish a living thing?” | SCORE: 0, 1, 2  
Target answer = yes  
(I) “yes” (score = 2)  
“no” (score = 0)  
+ REFER TO SCORING GUIDE #5 and #7 |
| (d) Assessor asks, “What makes you  | SCORE: 0, 1, 2, 3 |
| think that a fish is/is not a living thing? | Target answer = 3 scientifically acceptable answers
| + REFER TO SCORING GUIDE #6 and #7 |

3. Identification #3: *(photograph #3: strawberry)*

(a) Assessor shows the photograph of a strawberry and asks, **“What is this?”**

| SCORE: 0, 1, 2 |
| Target answer = strawberry/fruit |
| (I) “strawberry” (score = 2) |
| (II) “plant/fruit” (no score, continue with prompt – “How is this plant/fruit called?”) |
| (III) Any partially correct answer, for example, fruit (after prompt), berry, etc (score = 1) |

| + REFER TO SCORING GUIDE #7 |

(b) Assessor asks, **“Is a strawberry an animal, a plant, or something different?”**

| SCORE: 0, 1, 2 |
| Target answer = plant |
| (I) “plant” (score = 2) |
| “animal/ something different” (score = 0) |
| (II) “fruit” (no score, continue with prompt – “Is a fruit an animal, a plant, or something different?”) |

| + REFER TO SCORING GUIDE #3 and #7 |

**Note:** Still provide this question if child said “plant” in question 3(a)

(c) Assessor asks, **“Is a strawberry a living thing?”**

| SCORE: 0, 1, 2 |
| Target answer = yes |
| (I) “yes” (score = 2) |
| “no” (score = 0) |

| + REFER TO SCORING GUIDE #5 and #7 |

(d) Assessor asks, **“What makes you think that a strawberry is/is not a living thing?”**

| SCORE: 0, 1, 2, 3 |
| Target answer = 3 scientifically acceptable answers |
| + REFER TO SCORING GUIDE #6 and #7 |
| Identification #4: (photograph #4: clouds) | SCORE: 0, 1, 2  
Target answer = cloud(s)  
(I) “clouds/clouds in the sky” (score = 2)  
(II) “sky” (no score, continue with prompt – “What is this in the sky?” by pointing to the clouds).  
(III) Any partially correct answer, for example, fog, etc (score = 1)  
+ REFER TO SCORING GUIDE #7 |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Assessor shows the photograph of some clouds and asks, “What is this?”</td>
<td>---</td>
</tr>
</tbody>
</table>
| (b) Assessor asks, “Is a cloud an animal, a plant, or something different?” | SCORE: 0, 1, 2  
Target answer = something different  
(I) “something different” (score = 2)  
“animal/plant” (score = 0)  
+ REFER TO SCORING GUIDE #1 and #7 |
| (c) Assessor asks, “Is a cloud a living thing?” | SCORE: 0, 1, 2  
Target answer = no  
(I) “no” (score = 2)  
“yes/sometimes” (score = 0)  
+ REFER TO SCORING GUIDE #4 and #7 |
| (d) Assessor asks, “What makes you think that a cloud is/is not a living thing?” | SCORE: 0, 1, 2, 3  
Target answer = 3 scientifically acceptable answers  
+ REFER TO SCORING GUIDE #6 and #7 |
| Identification #5: (photograph #5: butterfly) | SCORE: 0, 1, 2  
Target answer = butterfly/insect  
(I) “butterfly” (score = 2)  
(II) “animal/insect” (no score, continue with prompt – “How is this animal/insect called?”)  
(III) Any partially correct answer, for example, insect (after prompt), bug, moth, fly, etc (score = 1)  
+ REFER TO SCORING GUIDE #7 |
<p>| (a) Assessor shows the photograph of a butterfly and asks, “What is this?” | --- |</p>
<table>
<thead>
<tr>
<th>Question</th>
<th>Score</th>
<th>Target Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>(b) Assessor asks, “Is a butterfly an animal, a plant, or something different?”</td>
<td>0, 1, 2</td>
<td>animal</td>
</tr>
<tr>
<td>Target answer = animal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(I) “animal” (score = 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“plant/something different” (score = 0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ REFER TO SCORING GUIDE #2 and #7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Note: Still provide this question if child said “animal” in question 5(a)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) Assessor asks, “Is a butterfly a living thing?”</td>
<td>0, 1, 2</td>
<td>yes</td>
</tr>
<tr>
<td>Target answer = yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(I) “yes” (score = 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“no” (score = 0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ REFER TO SCORING GUIDE #5 and #7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) Assessor asks, “What makes you think that a butterfly is/is not a living thing?”</td>
<td>0, 1, 2, 3</td>
<td>3 scientifically acceptable answers</td>
</tr>
<tr>
<td>Target answer = 3 scientifically acceptable answers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ REFER TO SCORING GUIDE #6 and #7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Identification #6: (photograph #6: flower)</td>
<td>0, 1, 2</td>
<td>flower/rose</td>
</tr>
<tr>
<td>(a) Assessor shows the photograph of a flower and asks, “What is this?”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target answer = flower/rose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(I) “flower/rose” (score = 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(II) “plant” (no score, continue with prompt – “How is this plant called?”)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(III) Any partially correct answer, for example, naming another type of flower (score = 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ REFER TO SCORING GUIDE #7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Assessor asks, “Is a flower an animal, a plant, or something different?”</td>
<td>0, 1, 2</td>
<td>plant</td>
</tr>
<tr>
<td>Target answer = plant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(I) “plant” (score = 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“animal/something different” (score = 0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question</td>
<td>Score</td>
<td>Target Answer</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>-------</td>
<td>---------------</td>
</tr>
<tr>
<td>(c) Assessor asks, “Is a flower a living thing?”</td>
<td>0, 1, 2</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) Assessor asks, “What makes you think that a flower is/is not a living thing?”</td>
<td>0, 1, 2, 3</td>
<td>3 scientifically acceptable answers</td>
</tr>
<tr>
<td>7. Identification #7: (photograph #7: grass)</td>
<td>0, 1, 2</td>
<td>grass</td>
</tr>
<tr>
<td>(a) Assessor shows the photograph of grass and asks, “What is this?”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Assessor asks, “Is grass an animal, a plant, or something different?”</td>
<td>0, 1, 2</td>
<td>plant</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) Assessor asks, “Is grass a living thing?”</td>
<td>0, 1, 2</td>
<td>yes</td>
</tr>
<tr>
<td>Question</td>
<td>Instruction</td>
<td>Expected Answer</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>(d) Assessor asks, “What makes you think that grass is/is not a living thing?”</td>
<td>SCORE: 0, 1, 2, 3. Target answer = 3 scientifically acceptable answers.</td>
<td></td>
</tr>
</tbody>
</table>
| (I) “yes” (score=2)  
“no” (score=0)                                                            | + REFER TO SCORING GUIDE #5 and #7                                           |                 |                |
| 8. Identification #8: (photograph #8: dog)                                | SCORE: 0, 1, 2. Target answer = dog/mammal                                  |                 |                |
| (a) Assessor shows the photograph of a dog and asks, “What is this?”    | (I) “dog/puppy” (score = 2)                                                 |                 |                |
| (II) “animal/mammal” (no score, continue with prompt – “How is this animal/mammal called?”) | (III) Any partially correct answer, for example, fox, wolf, etc (score = 1) |                 |                |
| (b) Assessor asks, “Is a dog an animal, a plant, or something different?” | SCORE: 0, 1, 2. Target answer = animal                                      |                 |                |
| (I) “animal” (score = 2)  
“plant/ something different” (score = 0)             | + REFER TO SCORING GUIDE #2 and #7                                           |                 |                |
| (c) Assessor asks, “Is a dog a living thing?”                             | SCORE: 0, 1, 2. Target answer = yes                                           |                 |                |
| (I) “yes” (score = 2)  
“no” (score = 0)                                                | + REFER TO SCORING GUIDE #5 and #7                                           |                 |                |
| (d) Assessor asks, “What makes you think that a dog is/is not a living thing?” | SCORE: 0, 1, 2, 3. Target answer = 3 scientifically acceptable answers |                 |                |
| 9. Identification #9: (photograph #9: airplane) | SCORE: 0, 1, 2  
Target answer = plane/airplane/aircraft  
(I) “plane/airplane/aircraft/jet/airbus” (score = 2)  
(II) “flying thing/machine” (no score, continue with prompt – “How is this flying thing/machine called?”)  
(III) Any partially correct answer, for example, flying thing/machine (after prompt), etc (score = 1)  
+ REFER TO SCORING GUIDE #6 and #7 |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Assessor shows the photograph of a plane and asks, “What is this?”</td>
<td></td>
</tr>
</tbody>
</table>
(b) Assessor asks, “Is an airplane an animal, a plant, or something different?” | SCORE: 0, 1, 2  
Target answer = something different  
(I) “something different” (score = 2)  
“animal/plant” (score = 0)  
+ REFER TO SCORING GUIDE #6 and #7 |
| (c) Assessor asks, “Is an airplane a living thing?” | SCORE: 0, 1, 2  
Target answer = no  
(I) “no” (score = 2)  
“yes/sometimes” (score = 0)  
+ REFER TO SCORING GUIDE #6 and #7 |
| (d) Assessor asks, “What makes you think that an airplane is/is not a living thing?” | SCORE: 0, 1, 2, 3  
Target answer = 3 scientifically acceptable answers  
+ REFER TO SCORING GUIDE #6 and #7 |
## Table A2

**Assessment Guide**

<table>
<thead>
<tr>
<th>Questions</th>
<th>What happens during assessment</th>
<th>How to proceed</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) to (d)</td>
<td>Child requests repetition of question</td>
<td>Repeat question to child</td>
</tr>
</tbody>
</table>
| (a) to (d) | Child says something that  
1) cannot be heard clearly  
2) is difficult to understand | Ask child to repeat answer |
| (a), (c), (d) only | Child does not provide any answer within 5-6 seconds after asking a question | Repeat question and allow an additional waiting time of 5 seconds before continuing assessment |
| (a) only | 1) Child cannot identify photograph after repeating question  
2) Child wrongly identifies photograph  
3) Child indicates verbally or non-verbally that he/she does not know the answer | Tell child the correct answer and then proceed with the assessment |
| (b) only | 1) Child does not provide any answer within 5-6 seconds after asking a question  
2) Child indicates verbally or non-verbally that he/she does not know the answer | Break down question as follows:  
1) Ask “Is a/an (item name) an animal?”  
2) Regardless of child’s answer or if there is no response within 5-6 seconds, or if child indicates verbally or non-verbally that he/she does not know the answer, ask “Is a/an (item name) a plant?”  
3) Again, regardless of child’s answer or if there is no response within 5-6 seconds, or if child indicates verbally or non-verbally that he/she does not know the answer, ask “Is a/an (item name) something different?”  
4) After this last sub question, allow a waiting time of 5-6 seconds (if there is no response), before continuing assessment. |
| (c) only | 1) Child does not provide any answer after repeating question  
2) Child indicates verbally or non-verbally that he/she does not know the answer | Skip part (d) of that set of questions |
| (d) only | 1) Child does not provide any answer after repeating question  
2) Child indicates verbally or non-verbally that he/she does not know the answer | Use the prompt “Can you tell me anything about it?” |
| (d) only | Child provides only one or two answers | Use the prompt “What else can you say about why a/an (item name) is /is not a living thing?”  
Provide no more than two prompts |
### Scoring Guide

<table>
<thead>
<tr>
<th>Number</th>
<th>Scoring Guide</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A</td>
<td>Pay attention to key terms children use for the part (d) questions — score responses irrespective of whether key terms are used in complete sentences or not.</td>
</tr>
</tbody>
</table>
| 1      | (a) Child is undecided/unsure but leaning toward correct  
> “Maybe/I think it’s something different” or words with similar meanings showing that the child believes it is something different (score = 1)  

(b) Child is undecided/unsure but leaning toward incorrect  
> “Maybe/I think it’s an animal/a plant” or words with similar meanings showing that the child believes it is an animal/a plant (score = 0)  

(c) Child replied “yes” to any two or all three sub questions (score = 0) |
| 2      | (a) Child is undecided/unsure but leaning toward correct  
> “Maybe/I think it’s an animal” or words with similar meanings showing that the child believes it is an animal (score = 1)  

(b) Child is undecided/unsure but leaning toward incorrect  
> “Maybe/I think it’s a plant/something different” or words with similar meanings showing that the child believes it is a plant/something different (score = 0)  

(c) Child replied “yes” to any two or all three sub questions (score = 0) |
| 3      | (a) Child is undecided/unsure but leaning toward correct  
> “Maybe/I think it’s a plant” or words with similar meanings showing that the child believes it is a plant ” (score = 1)  

(b) Child is undecided/unsure but leaning toward incorrect  
> “Maybe/I think it’s an animal/something different” or words with similar meanings showing that the child believes it is an animal/something different (score = 0)  

(c) Child replied “yes” to any two or all three sub questions (score = 0) |
| 4      | (a) Child is undecided/unsure but leaning toward correct  
> “Maybe/I think it’s not living” or words with similar meanings showing that the child believes it is not a living thing (score = 1)  

(b) Child is undecided/unsure but leaning toward incorrect  
> “Maybe/I think it’s living” or words with similar meanings showing that the child believes it is a living thing (score = 0) |
| 5      | (a) “Sometimes” followed by no explanation (no score, continue with prompt “When is it living?”)  
- Wrong or incorrect explanation, for example, “when the fish is grey”; “when the flower is pink” (score = 1)  
- Correctly mentioned circumstances regarding when item is/is not living, for example, “when the fish swims”; “when the flower grows” (score = 2) |
### (b) Child is undecided/unsure but leaning toward correct

“Maybe/I think it’s living” or words with similar meanings showing that the child believes it is a living thing” (score = 1)

### (c) Child is undecided/unsure but leaning toward incorrect

“Maybe/I think it’s not living” or words with similar meanings showing that the child believes it is not a living thing (score = 0)

<table>
<thead>
<tr>
<th>6</th>
<th>(a) Score part (d) of each set of questions relative to child’s answer in the corresponding part (c) question. See examples in table A4.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(b) Any scientifically acceptable answer (score = 1)</td>
</tr>
<tr>
<td></td>
<td>(i) For living things: answer should reflect characteristics of living things such as growth, movement, reproduction, nutrition, excretion, respiration (use and release of energy), and sensitivity/responsiveness</td>
</tr>
<tr>
<td></td>
<td>(ii) For non-living things: answer should reflect the opposite of the characteristics of living things mentioned in part b(i) above.</td>
</tr>
<tr>
<td></td>
<td>See brief description of characteristics of living things and corresponding examples in table A5.</td>
</tr>
<tr>
<td></td>
<td>(c) Answer is scientifically unacceptable or is unrelated to question (score = 0)</td>
</tr>
<tr>
<td></td>
<td>(d) Child provides three or more scientifically acceptable answers (score = 3)</td>
</tr>
</tbody>
</table>

| 7 | There is no response from child after repeating question (score = 0)                                                                                                                   |
|   | Answer is incorrect and/or unrelated to question (score = 0)                                                                                                                         |
|   | Child indicates verbally or non-verbally that he/she does not know the answer (score = 0)                                                                                            |
Table A4

Examples and Explanations for Particular Scores for Part (d) Questions

<table>
<thead>
<tr>
<th>Item</th>
<th>Answer to part (c)</th>
<th>Answer to part (d)</th>
<th>Score</th>
<th>Brief explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>fish</td>
<td>living</td>
<td>It moves</td>
<td>1</td>
<td>Correctly attributing movement as characteristic of living things</td>
</tr>
<tr>
<td>fish</td>
<td>not living</td>
<td>It moves</td>
<td>0</td>
<td>Incorrectly attributing movement as characteristic of non-living things by identifying fish as a non-living thing</td>
</tr>
<tr>
<td>fish</td>
<td>living</td>
<td>It does not move</td>
<td>0</td>
<td>Incorrectly attributing absence of movement as characteristic of living things</td>
</tr>
<tr>
<td>fish</td>
<td>not living</td>
<td>It does not move</td>
<td>0</td>
<td>Incorrectly attributing absence of movement to a living thing</td>
</tr>
<tr>
<td>fish</td>
<td>living/non-living</td>
<td>I have a fish at home</td>
<td>0</td>
<td>Scientifically unacceptable characteristic of living and non-living things</td>
</tr>
<tr>
<td>snowman</td>
<td>not living</td>
<td>It does not grow</td>
<td>1</td>
<td>Correctly attributing absence of growth as characteristic of non-living things</td>
</tr>
<tr>
<td>snowman</td>
<td>living</td>
<td>It does not grow</td>
<td>0</td>
<td>Incorrectly attributing absence of growth as characteristic of living things by identifying snowman as a living thing</td>
</tr>
<tr>
<td>snowman</td>
<td>not living</td>
<td>It grows</td>
<td>0</td>
<td>Incorrectly attributing growth as characteristic of non-living things</td>
</tr>
<tr>
<td>snowman</td>
<td>living</td>
<td>It grows</td>
<td>0</td>
<td>Incorrectly attributing growth to a non-living thing</td>
</tr>
<tr>
<td>snowman</td>
<td>living/not living</td>
<td>It is white</td>
<td>0</td>
<td>Scientifically unacceptable characteristic of living and non-living things</td>
</tr>
<tr>
<td>clouds</td>
<td>living/not living</td>
<td>It moves</td>
<td>0</td>
<td>Clouds do not move by themselves and also do not possess all of the other scientifically acceptable characteristics of living things</td>
</tr>
<tr>
<td>airplane</td>
<td>living/not living</td>
<td>It moves/flies</td>
<td>0</td>
<td>Airplanes do not move by themselves and also do not possess all of the other scientifically acceptable characteristics of living things</td>
</tr>
<tr>
<td>clouds/airplane</td>
<td>not living</td>
<td>It does not move</td>
<td>1</td>
<td>Moving independently is a scientifically acceptable characteristic of living things, and clouds and airplanes do not move by themselves</td>
</tr>
</tbody>
</table>
Table A5

**Brief Description of Characteristics of Living Things and Examples**

<table>
<thead>
<tr>
<th>Characteristics of living things</th>
<th>Brief description</th>
<th>Examples</th>
</tr>
</thead>
</table>
| **Growth**                       | Increase in the size and complexity of an organism or part of an organism, usually due to an increase in the number of cells | 1) Fish egg turning into fish  
2) Seed turning into flowering plant |
| **Movement**                     | Act of changing physical location or position by oneself | 1) Dog walking/running  
2) Butterfly flying  
3) Fish swimming  
4) Bending of flowering plant’s shoot toward light  
5) Moving of plant roots toward underground water |
| **Reproduction**                 | Natural process among organisms resulting in the production of new individuals of the same kind | 1) Fish/butterfly laying eggs  
2) Dog giving birth to young ones  
3) Flowering plants producing seeds |
| **Nutrition**                    | Process by which an organism obtains food which is used to provide materials and energy to assist in life maintaining activities | 1) Dog eating food and drinking water  
2) Grass roots taking water and mineral ions from soil |
| **Excretion**                    | Elimination of waste matter from the body | 1) Dog peeing  
2) Grass expelling oxygen |

*Note: Excretion and egestion are different processes. Excretion is the removal of waste product of metabolism from the body (e.g. urine) whereas egestion is the removal of undigested food from the gut (e.g. feces).*

| **Respiration**                  | Use and production of energy | 1) Fish using energy from food to swim  
2) Strawberry plant using energy from sunlight to produce sugar to be used as energy |

*Note: respiration and breathing are different processes. Breathing is the physical act of inhaling and exhaling whereas respiration encompasses several metabolic processes.*

| **Sensitivity/Responsiveness**   | Ability to respond to stimuli in the external environment | 1) Dog seeing/hearing/touching/tasting/smelling  
2) Leaves of flowering plants moving toward the sun  
3) Roots of flowering plants growing down to search for water and minerals |
Appendix B

Photographs Used in the LSA

Figure B1. Fish
Figure B2. Butterfly
Figure B3. Dog
Figure B4. Strawberry
Figure B5. Grass
Figure B6. Flower
Figure B7. Snowman
Figure B8. Airplane
Figure B9. Clouds
Appendix C

Table C1

Table of Specifications for the Life Sciences Assessment

<table>
<thead>
<tr>
<th>Objectives</th>
<th>Content</th>
<th>Number of items</th>
<th>Item number</th>
<th>Percentage of test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify living things (animals and plants) and non-living things from photographs by providing their common names</td>
<td>9 photographed items: 3 animals – fish, butterfly, and dog; 3 plants – strawberry, grass, and flower; 3 non-living things – snowman, clouds, and airplane</td>
<td>9</td>
<td>1a, 2a, 3a, 4a, 5a, 6a, 7a, 8a, 9a</td>
<td>25</td>
</tr>
<tr>
<td>Distinguish photographed items as animal, plant, or something different</td>
<td>9 photographed items: 3 animals – fish, butterfly, and dog; 3 plants – strawberry, grass, and flower; 3 non-living things – snowman, clouds, and airplane</td>
<td>9</td>
<td>1b, 2b, 3b, 4b, 5b, 6b, 7b, 8b, 9b</td>
<td>25</td>
</tr>
<tr>
<td>Distinguish whether photographed items were living or non-living</td>
<td>9 photographed items: 3 animals – fish, butterfly, and dog; 3 plants – strawberry, grass, and flower; 3 non-living things – snowman, clouds, and airplane</td>
<td>9</td>
<td>1c, 2c, 3c, 4c, 5c, 6c, 7c, 8c, 9c</td>
<td>25</td>
</tr>
<tr>
<td>Provide scientific explanation regarding why photographed items were considered as living or non-living</td>
<td>9 photographed items: 3 animals – fish, butterfly, and dog; 3 plants – strawberry, grass, and flower; 3 non-living things – snowman, clouds, and airplane</td>
<td>9</td>
<td>1d, 2d, 3d, 4d, 5d, 6d, 7d, 8d, 9d</td>
<td>25</td>
</tr>
</tbody>
</table>
### Appendix D

**Tables**

**Table D1**

*Descriptive Statistics for the Scales of the LSA and the Total Measure*

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naming</td>
<td>17.15</td>
<td>1.354</td>
<td>-1.787</td>
<td>3.045</td>
</tr>
<tr>
<td>Classifying</td>
<td>11.93</td>
<td>4.191</td>
<td>-.362</td>
<td>-.980</td>
</tr>
<tr>
<td>Living</td>
<td>10.67</td>
<td>2.957</td>
<td>-.068</td>
<td>.379</td>
</tr>
<tr>
<td>Explaining</td>
<td>1.56</td>
<td>1.914</td>
<td>1.236</td>
<td>.581</td>
</tr>
<tr>
<td>Total LSA</td>
<td>41.23</td>
<td>8.022</td>
<td>-.208</td>
<td>-.881</td>
</tr>
</tbody>
</table>

*Note.* Classifying = classification: animal, plant, or something different (CL_apsd); Living = classification: living or non-living (CL_lnl); Explaining = scientific explanation: living or non-living (SE_lnl); LSA = Life Sciences Assessment
Table D2

*Kolmogorov-Smirnov Values for the Scales of the LSA and the Total Measure*

<table>
<thead>
<tr>
<th>Scale</th>
<th>Z Statistic</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naming</td>
<td>3.977</td>
<td>.000</td>
</tr>
<tr>
<td>Classifying</td>
<td>1.841</td>
<td>.002</td>
</tr>
<tr>
<td>Living</td>
<td>1.787</td>
<td>.003</td>
</tr>
<tr>
<td>Explaining</td>
<td>2.908</td>
<td>.000</td>
</tr>
<tr>
<td>Total LSA</td>
<td>1.203</td>
<td>.111</td>
</tr>
</tbody>
</table>

*Note.* Classifying = classification: animal, plant, or something different (CL_apsd); Living = classification: living or non-living (CL_lnl); Explaining = scientific explanation: living or non-living (SE_lnl); LSA = Life Sciences Assessment
Table D3

*Correlations Among Scales*

<table>
<thead>
<tr>
<th>Scale</th>
<th>Naming</th>
<th>Classifying</th>
<th>Living</th>
<th>Explaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naming</td>
<td>1</td>
<td>.535**</td>
<td>.195*</td>
<td>.285**</td>
</tr>
<tr>
<td>Classifying</td>
<td>1</td>
<td>.422**</td>
<td>.449**</td>
<td></td>
</tr>
<tr>
<td>Living</td>
<td></td>
<td></td>
<td>1</td>
<td>.476**</td>
</tr>
<tr>
<td>Explaining</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

*Note.* Classifying = classification: animal, plant, or something different (CL_apsd); Living = classification: living or non-living (CL_lnl); Explaining = scientific explanation: living or non-living (SE_lnl); ** = Correlation is significant at the 0.01 level (2-tailed); * = Correlation is significant at the 0.05 level (2-tailed)
Table D4

*Correlations Between Age in Months and Scores on the LSA*

<table>
<thead>
<tr>
<th>Scale</th>
<th>r</th>
<th>Sig</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naming</td>
<td>.122</td>
<td>.176</td>
<td>.015</td>
</tr>
<tr>
<td>Classifying</td>
<td>.297*</td>
<td>.001</td>
<td>.088</td>
</tr>
<tr>
<td>Living</td>
<td>.166</td>
<td>.065</td>
<td>.028</td>
</tr>
<tr>
<td>Explaining</td>
<td>.281*</td>
<td>.003</td>
<td>.079</td>
</tr>
<tr>
<td>Total LSA</td>
<td>.306*</td>
<td>.001</td>
<td>.093</td>
</tr>
</tbody>
</table>

*Note.* Classifying = classification: animal, plant, or something different (CL_apsd); Living = classification: living or non-living (CL_lnl); Explaining = scientific explanation: living or non-living (SE_lnl); LSA = Life Sciences Assessment; ** = Correlation is significant at the 0.01 level (2-tailed)
Table D5

*Test for Univariate Outliers on the Scales of the LSA*

<table>
<thead>
<tr>
<th>Scale</th>
<th>Mean</th>
<th>Standard Deviation (SD)</th>
<th>Mean ± 4 x SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naming</td>
<td>17.15</td>
<td>1.354</td>
<td>11.734, 22.566</td>
</tr>
<tr>
<td>Classifying</td>
<td>11.93</td>
<td>4.191</td>
<td>-4.834, 28.694</td>
</tr>
<tr>
<td>Living</td>
<td>10.67</td>
<td>2.957</td>
<td>-1.158, 22.498</td>
</tr>
<tr>
<td>Explaining</td>
<td>1.56</td>
<td>1.914</td>
<td>-6.096, 9.216</td>
</tr>
</tbody>
</table>

*Note.* Classifying = classification: animal, plant, or something different (CL_apsd); Living = classification: living or non-living (CL_lnl); Explaining = scientific explanation: living or non-living (SE_lnl)
Appendix E

Data Figures

Figure E1. Histograms representing total scores and frequencies on the LSA
Figure E2. Scatterplots comparing age in months to scores on the LSA
Figure E3. Hypothesized CFA model for the LSA.
Figure E4. Standardized factor loadings of the scales of the LSA