2017

Validation Study of the Science Literacy Assessment: A Measure to Assess Middle School Students' Attitudes Toward Science and Ability to Think Scientifically

Tammy McKeown
Virginia Commonwealth University

VALIDATION STUDY OF THE SCIENCE LITERACY ASSESSMENT: A MEASURE TO ASSESS MIDDLE SCHOOL STUDENTS' ATTITUDES TOWARD SCIENCE AND ABILITY TO THINK SCIENTIFICALLY

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy at Virginia Commonwealth University

by

Tammy R. McKeown

Bachelor of Science, Virginia Commonwealth University, 2004

Master of Science, Virginia Commonwealth University, 2006

Director: Lisa M. Abrams, Ph.D.

Associate Professor, Foundations of Education

Virginia Commonwealth University

Richmond, Virginia

September, 2017

Acknowledgement

As I reflect on this journey, I am overwhelmed with gratitude for the amazing people who supported me through this process. Without the love and encouragement of my family and friends, this dissertation would never have come to fruition. I am eternally grateful for those who kept me motivated, provided words of encouragement, and believed in me when I couldn't believe in myself. I consider myself to be truly blessed, and I thank my Creator for placing each of these special people into my life, just when I needed them.

I would also like to thank my committee members for their guidance, support, and thoughtful feedback: Dr. Lisa Abrams (chair), Dr. Patricia Slattum, Dr. Jacqueline McDonnough, and Dr. Angie Wetzel. I am very appreciative to each of you for the time and effort you devoted to assisting me, for pushing me to do my best work, and for remaining patient as I struggled to find my way.

Special thanks to those who participated in the study, particularly the students. You brought many smiles to my face as I watched you being silly in class, admired the pictures and doodles you drew on your survey packets, and reminded me of the energy and enthusiasm that often radiates from those who are young and curious.

Lastly, I would like to acknowledge the grant that made this research possible. The research reported here was supported in part by the National Institutes of Health, Office of the Director, Science Education Partnership Award through Grant R25 OD010984-05 to the Virginia Commonwealth University. The opinions expressed are those of the author and do not represent the views of NIH.

Table of Contents

List of Tables viii

List of Figures ix

Introduction 1

Statement of the Problem 3

Overview of the Literature 4

Justifications for Promoting Science Literacy 4

Conceptualizations of Science Literacy 5

Science Education Reform Movement in the US 6

Test-Based Accountability Policy and Science Education 8

Measuring Scientific Literacy 9

Calls to Measure Scientific Literacy as a Way of Thinking 10

Student Attitudes Toward Science 11

Measuring Student Attitudes Toward Science 12

Rationale and Purpose of the Study 13

Research Questions 14

Definition of Terms 15

Review of the Literature 16

Method for Review of the Literature 16

Justifications for Promoting Science Literacy 17

Benefits of Increased Science Literacy at the Micro Level 18

Benefits of Increased Science Literacy at the Macro Level 19

Conceptualizations of Science Literacy 21

Analyzing Educational Literature to Synthesize a Definition of Science Literacy 22

Types and Levels of Science Literacy Based on the Assumed Needs of Learners 25

Conceptualizations of Science Literacy Originating from Scientists 26

Defining Science Literacy Within the Context of Everyday Life 27

Science Education Reform in the United States 31

Wave I: Science Education Reform following World War II 32

Wave II: Science Education Reform Since the 1980s 34

NSTA Project on Scope, Sequence, and Coordination of Secondary School Science 35

AAAS Project 2061 36

NRC National Science Education Standards 36

Test-Based Accountability and Science Education 40

No Child Left Behind 40

Every Student Succeeds Act 41

Science Education Policy and Science Instruction 41

Measurement of Student Scientific Literacy 43

Methodological Approaches 44

Calls to Measure Student Scientific Literacy as a Way of Thinking 46

Student Attitudes Toward Science 50

Self-Efficacy 52

Task Value 53

Student Epistemological Beliefs 55

Measuring Student Attitudes Toward Science 58

Summary 59

Methods 61

Research Design 63

Recruitment 63

Participants 64

Original Data Collection 66

Measurement 66

Assessment of Student Science Literacy 66

Assessment of Student Attitudes Toward Science 67

Assessment of Student Interest in Science Careers 68

Assessment of Student Science General Knowledge 68

Prescreening 72

Completeness 73

Outliers 73

Characteristics of the Sample 74

Delimitations 71

Summary 71

Findings 72

Research Question 1 82

Exploratory Factor Analysis of the SLA-MB 82

Exploratory Factor Analysis of the SLA-MD....................79

Research Question 2 88

Exploratory Factor Analysis of the SLA-MB....................88

Exploratory Factor Analysis of the SLA-D 92

Research Question 3 93

Pearson Product-moment Correlations 93

Research Question 4 95

Bivariate Linear Regression - Middle School Sample....................93

Bivariate Linear Regression - High School Sample 93

Discussion....................96

Interpretation of Key Findings....................65

Validity Evidence Based on Internal Structure....................96

Concurrent Criterion-Related Validity 100

Predictive Criterion-Related Validity 101

Implications for Research and Practice....................102

Study Limitations....................104

Recommendations for Future Research....................105

Conclusion 106

References....................108

Appendices....................123

A SLA-D....................123

B SLA-MB....................128

C Modified Attitudes Toward Science Inventory....................130

D STEM Career Interest Survey 132

E Middle School General Science Knowledge Survey..133

F High School General Science Knowledge Survey 143

Tables

Table 1: Research Questions, Sources of Validity Evidence, & Analytic Procedures 65

Table 2: Characteristics of Students Completing Survey Packets 67

Table 3: Mapping of Items for High School General Science Knowledge Survey 72

Table 4: Mapping of Items for Middle School General Science Knowledge Survey 73

Table 5: Description of Missing Responses and Cases Deleted Due to Missing Data 77

Table 6: Description of Variables Examined for Each Construct...80

Table 7: Descriptive Data for SLA-Motivation and Beliefs Items 81

Table 8: Mean and Standard Deviation of MATSI Items/Scales 84

Table 9: Descriptive Data for STEM-CIS 86

Table 10: Three Factor solution for the SLA-MB Middle School Sample 91

Table 11: Three Factor solution for the SLA-MB High School Sample 97

Table 12: Correlations Between Student Scores on the SLA-MB, STEM-CIS, and mATSI 101

Figures

Figure 1: Initial Unrotated Solution for SLA-MB Middle School Sample 89

Figure 2: Path Diagram of Three Factor Solution for the SLA-MB Middle School Sample 92

Figure 3: Initial Unrotated Solution for SLA-D Middle School Sample.......................................94

Figure 4: Initial Unrotated Solution for SLA-MB High School Sample 95

Figure 3: Path Diagram of Three Factor Solution for the SLA-MB High School Sample...........98

Abstract

VALIDATION STUDY OF THE SCIENCE LITERACY ASSESSMENT: A MEASURE TO ASSESS MIDDLE SCHOOL STUDENTS' ATTITUDES TOWARD SCIENCE AND ABILITY TO THINK SCIENTIFICALLY

By Tammy R. McKeown, MS

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University, School of Education

Dissertation Chair: Lisa M. Abrams, Ph.D.

Associate Professor, Foundations of Education

Virginia Commonwealth University, 2017

The present study investigated validity evidence for the Science Literacy Assessment, an instrument designed to assess middle school students' ability to think scientifically as well as their motivation and beliefs about science (Fives, Huebner, Birnbaum, & Nicolich, 2014). Specifically, three sources of evidence were considered; internal structure, concurrent criterion-related, and predictive criterion-related. Exploratory factor analysis was utilized to examine the underlying factor structure of each of the instrument's two components, motivation and beliefs related to science and demonstrated scientific literacy. Pearson product-moment correlations were calculated to determine the relationship between scores on the motivation and belief component of the Science Literacy Assessment and two instruments widely used to assess students' attitudes toward science, the Modified Attitudes Toward Science Inventory (Weinburg

& Steele, 2000), and the STEM Career Interest Survey (Kier, Blanchard, Osborne, & Albert, 2014). Finally, the extent to which scores on the Science Literacy Assessment predict scores on a general science knowledge instrument was assessed with the use of bivariate linear regression. Results suggested that, for the middle school student sample, the Science Literacy Assessment has appropriate psychometric properties for use with middle school students. Due to an insufficient high school sample size, validity evidence for this group was inconclusive.

Chapter 1

Introduction

The launch of Sputnik 1 in 1957 introduced an era of rapid advancement in science and technology that ignited the imagination of much of the world (Graber, 2007). During the eight-day mission of Apollo 11 in 1969, television networks broadcast 280 hours of coverage to over 500 million people around the globe (NASA, 2015). Along with the rapid developments during this period in history, appropriately termed the Space Age, came discussions of the public's ability to understanding science. In 1958, Paul Hurd coined the term "science literacy" stating, "The American people, sparked by a Sputnik, and almost as a single voice, have inquired whether their children are receiving the kind of education that will enable them to cope with a society of expanding scientific and technological developments" (Hurd, 1958, pp. 13-14). Hurd's comments marked the beginning of a shift in mainstream science educational reform, moving from a system which focused on producing scientifically knowledgeable individuals who were bound for careers as scientists and engineers to one that aimed to achieve science literacy for the general population (Wenning, 2006).

The reconceptualization of science education that began in the 1950s has continued to be driven by advances in technology and science as they transform our word at a rapid pace. Given the degree to which modern cultures rely on these advances, it has become widely accepted that science literacy is an important educational and societal goal (NAS, 2007). For those who live in industrialized countries, science and technology are an inescapable component of every-day life. Americans encounter science in their roles as consumers, employees, and citizens. They

compete for jobs in technology-driven sectors that didn't exist a decade ago. They vote for political candidates with differing views on climate change, energy programs, and stem cell research. And, they rely on highly sophisticated technological devices. Science has also changed the way we communicate, our methods of transportation, our food and clothing, and even the length and quality of life. Along with these changes come discussions of philosophy and moral values as science has even given us the ability to destroy ourselves (Impey et al., 2011).

Despite the important role science plays within modern society, achieving a scientifically literate populace has proven to be a difficult task in the United States. Students in the U.S. have fallen behind in science and mathematics achievement when compared to their international peers (National Science Board, 2006). Research over the past four decades has found wide-spread misconceptions of basic science concepts among K-16 students (Baker, 2004; Wandersee, Mintzes & Novak, 1994). The US National Assessment of Educational Progress (NAEP) reported in the 2011 issue of The Nation's Report Card that only 65 percent of eighth-grade students performed at or above the Basic proficiency level – the lowest of NAEP's three achievement levels and indicates a partial mastery of fundamental skills. Only 32 percent performed at or above Proficient, and two percent of students performed at the Advanced level (National Center for Educational Statistics, 2012). While the United States spends more money per student and has more highly educated parents than most countries, these advantages have not translated into higher student performance. The most recent Programme for International Assessment (PISA), a triennial survey of over 510,000 15-year-old students in 65 countries around the world, indicated that students in the United States ranked slightly below average in science literacy (OECD, 2014).

The low performance of youth in the United States on science assessments has led to calls for rethinking the content and pedagogy of science education. Researchers have suggested that achievement in science is influenced by factors such as: student abilities, perceptions and attitudes, socioeconomic status, parental expectations, peer influences, and school-related variables (Singh, Granville, & Dika, 2002). Many of these components, such as family and home-related factors, are difficult to change and are outside the control of educators. However, there are several factors that can be influenced by educational interventions such as students' (a) academic engagement, (b) perceptions and attitudes towards science, and (c) knowledge of the role of science achievement in future career opportunities (Singh, Granville, & Dika, 2002). However, developing educational interventions to address the low science literacy of this nation's youth has many challenges. As the concept of science literacy has evolved, a number of interpretations and definitions have been proposed. Because of this, there is no universal definition of science literacy which has implications for both research and practice (DeBoer, 1991). Further complicating the development of academic interventions that promote science literacy as a way of knowing are the methods commonly used to measure the concept. For example, existing measures of science literacy are largely dependent on discipline specific content (Jenkins, 2003). Additionally, Lambert (2006) claimed that several state and national standardized tests attempt to measure science literacy, but the scope of these tests is so broad that teachers must engage in surface level coverage of a wide range of topics, which results is an emphasis on content coverage, rather than a deep focus on essential concepts.

Statement of the Problem

A common concern expressed in reports on science education is that the content and pedagogical approaches within K-12 classrooms are not aligned with the interests of the majority of students or the needs of society (Hofstein, Eilks, & Bybee, 2011). For example, the general public rarely interacts with scientific concepts that are decontextualized. Rather, most individuals encounter science within the complex social and political climate which is deeply embedded within our culture (Sinatra, Kienhues, & Hoffer, 2014). Because of this fact, many have called for science education to focus on science as inquiry rather than memorization of isolated and field-specific facts. In order to achieve this goal, educators will need to provide students with a curriculum that develops science literacy as a way of knowing. However, existing measures of science literacy are largely dependent on field-specific content, isolated from the process of science inquiry. In contrast, assessment of science literacy would benefit from an emphasis on those aspects of science that transcend discipline specific content, focus on the process of science, and reflect scientific training (Jenkins, 2003). Additionally, most current measures target students at the secondary or university level, and they do not include assessment of students' motivation and beliefs about science (Fives, Huebner, Birnbaum, & Nicolich, 2014). The inclusion of a measure to examine not only science literacy, but also students' attitudes toward science would advance science pedagogy as research has linked student motivation and beliefs to science achievement (Britner & Pajares, 2006; Chen & Pajares, 2010).

Overview of the Literature

The following brief review of the literature includes justifications for promoting science literacy as well as descriptions of various conceptualizations of science literacy. Science education reform efforts within the United States are also described and literature regarding the

measurement of science literacy is discussed. Finally, the link between student attitudes and achievement is addressed, as are efforts to measure student attitudes towards science.

Justifications for Promoting Science literacy

There are an abundance of reasons for advancing the science literacy of individuals as well as the general public. Arguments justifying science literacy are used in two different ways. The first refers to students and/or adults, in terms of the characteristics of the scientifically literate person. The second refers to system wide (national, regional, local) curriculum policies (Roberts 2007). Laetsch (1987) categorized common arguments for science literacy as: (a) science literacy allows for better political decisions, (b) science literacy facilitates improved individual behaviors, and (c) science literacy helps to create a more ethical world. Similarly, the National Research Council (NRC) emphasizes that science should be a part of basic education, because every individual needs some knowledge of science (NRC, 2007). According to the NRC, school science education should promote science literacy because:

1. Science is a significant part of human culture and represents one of the pinnacles of human thinking capacity;
2. It provides a laboratory of common experience for development of language, logic, and problem-solving skills in the classroom';
3. A democracy demands that its citizens make personal and community decisions about issues in which scientific information plays a fundamental role, and they hence need a knowledge of science as well as an understanding of scientific methodology;
4. For some students, it will become a lifelong vocation or avocation; and
5. The nation is dependent on the technical and scientific abilities of its citizens for its economic competitiveness and national needs (NRC, 2007, p. 34).

Conceptualizations of Science literacy

While there has been a great deal written about science literacy, there is no consensus on its meaning or essential components. Roberts (2007) suggests the large amount of literature on the topic can be more easily understood by considering five of the approaches, or conceptual methodologies, that authors have used in writing about the construct. One line in the literature is historical, rooted in the discussions of professional science educators who have tried to synthesize and make sense of the multiple definitions of science literacy appearing in the literature between 1960 and 1980. Another concentrates on "types" and "levels" of science literacy in term of justification arguments based on the assumed needs of learners. A third line conceptualizes meaning for science literacy by concentrating on different interpretations of what it means to be literate, and a fourth addresses science literacy by focusing on science and scientists. Finally, there is the approach that examines situations or contexts in which aspects of science are presumed to be of value in students' everyday lives.

Science Education Reform Movement in the US

Science became part of the school curriculum during the 19th century, largely because of the efforts of scientists themselves who argued that students would gain the highest level of education through the inductive process of observing the natural world and then making conclusions. This level of scientific knowledge was said to further benefit students by enabling them to participate more fully in a democratic society (DeBoer, 2000). Early in the 20th century, science education was justified by writers such as John Dewey on the basis of its relevance to contemporary life and its contributions to a shared understanding of the world by members of society. In 1902, Dewey argued that teaching theory should be more aligned with desired outcomes and that the best way to help students learn is through experiential learning. He stated

that in school “Facts are torn away from their original place in experience and rearranged with reference to some general principal” (p.6).

From the mid-1940s through the 1960s, educational reform was guided by the work of psychologists such as Jean Piaget’s theory of cognitive development and Kurt Lewin’s insights on child development. Reform discussions were founded on questions such as the intellectual ability of children at various stages of development and how this might impact pedagogy. During this period, the major motive for reform was to produce more scientists and the instructional methods of most science educators focused on delivering content knowledge through the use of drill and practice, common strategies at the time.

By the early 1980s, science education reform began placing substantial attention on a broader form of science literacy, with reform efforts driving major changes aimed at increasing science literacy for all children, not just those interested in a scientific career (Wenning, 2006). This shift in focus was sparked by a 1983 report from the National Commission on Excellence in Education, *A Nation at Risk: The Imperative for Educational Reform*. The report pointed to the poor academic performance of students in the United States as the cause of a declining position in world markets and suggested the solution was to create more rigorous curricula for English, mathematics, science, social studies, computer science, and foreign languages (DeBoer, 2000). The heightened sense of urgency caused by the report resulted in major, national-level educational reforms aimed at increasing the degree of science literacy among students, and eventually the general public (Wenning, 2006). To accomplish this goal, the National Science Education Standards were established by the National Research Council (NRC) in 1996. The NRC went on to further refine standards based reform, releasing *A Framework for K-12 Science Education Standards* in 2011. The framework presented a broad description of the content all

students were expected to learn by the completion of high school (NRC, 2011). The framework was then utilized to produce K-12 science standards, designed to be rich in content and practice, providing all students an internationally benchmarked science education. Development of the *Next Generation Science Standards* (NGSS, 2013) was a collaborative process including a consortium of 26 states, the National Science Teachers Association, the American Association for the Advancement of Science, the National Research Council, and Achieve, a nonprofit organization. The standards include scientific practices that promote critical thinking and communication skills by utilizing three dimensions woven into lessons at all grade levels. The first is disciplinary core ideas, which are central to each field of science and guide learners in observing, thinking, explaining, and problem solving. The second is crosscutting concepts, which serve as a mechanism for connecting ideas across science disciplines. The third dimension, practices, describe and teach students the behaviors utilized by scientists. Practices focus on teaching students to generate questions, investigate and analyze data, construct models, and refine explanations. The standards also outline "performance expectations," defining what students must be able to do in order to demonstrate competency (Krajcik, Codere, Dahsah, Bayer, & Mum, 2014).

Test-Based Accountability Policy and Science Education

Given the vital role of science education in the United States, it is important to understand how testing policy influences the teaching and learning of science within schools. Testing policies operate concurrently with instructional reform efforts and can influence the nature and focus of the content skills at the focus of classroom teaching. In 1965, the Elementary and Secondary Education Act (ESEA) set guidelines to evaluate how well education systems utilized federal funding and began using quantitative measurements to evaluate

educational quality and student achievement. President Nixon further promoted the idea of attaching quantitative measures to educational quality when he addressed Congress in 1970, emphasizing the need for “measurable” standards (Sirotnik, 2004, p. 150). Accountability legislation had passed in 27 states by 1973 (DeNovellis & Lewis, 1974) and the practice of holding schools accountable through the use of educational testing continued to expand through the 1990s (Anderson, 2012). The reauthorization of ESEA in 1988 required yearly examination of educational progress through student testing in schools and districts receiving Title I funding, and ESEA 1994 linked content standards and student testing to accountability (Penfield & Lee, 2009). The term “high stakes” testing came to be common place when states imposed sanctions on schools if they did not meet benchmarks that represented success. In 2001, the No Child Left Behind Act (NCLB) renewed ESEA and further increased the focus on test-based accountability, requiring all states to administer yearly assessments and imposing sanctions on schools that did not meet set standards (Anderson, 2012).

Early in the testing movement the symbolic nature of test scores was recognized. Airasian (1987) contends that the general public and policymakers place a great deal of trust in standardized tests and are in favor of implementing these tests due to several appealing factors. Airasian believes most people perceive standardized tests to be ‘fair’ because all students are required to demonstrate proficiency on identical tests, ‘scientific’ because the tests produce a numerical score, and ‘objective’ because decisions based on the numerical scores are not typically perceived as being influenced by personal biases.

Measuring Science literacy

The assessment of students’ science literacy is an outgrowth of statewide accountability systems. Even though policy makers and educators proclaim the importance of science

education, a system for the comprehensive measurement of students' science literacy remains elusive. Although there are currently several measures of science literacy (e.g., Bybee, 2009; OECD, 2006; Wenning, 2006, 2007) most instruments measure content knowledge of one or more specific science field/disciplines, are developed for students at the secondary or post-secondary level, and fail to assess students' motivation and beliefs about science (Fives, Huebner, Birnbaum, & Nicoloch, 2014). This lack of an effective measure has been attributed, in part, to the wide range of definitions and conceptualizations of science literacy (Wenning, 2006), differing methodological approaches to measuring the construct (Laugksch, 2000), and an over-emphasis on discipline specific content necessitated by the pressure of standardized testing (Rudolph, 2014).

Calls to Measure Student Science literacy as a Way of Thinking. Advocacy for providing students with a more complex and multidimensional understanding of science is not a new phenomenon. In 1910, John Dewey argued that science education should promote what he referred to as "the scientific habit of mind" providing students with the ability to engage with scientific reasoning within everyday life (Dewey, 1910, p. 126). Following Dewey's appeal, others have promoted the belief that science literacy should be a life-long and evolving state rather than the achievement of a goal (Liu, 2009).

Expanding on this perspective, Feinstein, Allen, & Jenkins (2013) asserted that science education should help students identify and interpret scientific information that is needed within their everyday lives. They asserted that the ability to evaluate the credibility of information based on evidence should be included in the process of science instruction. This reconsideration of science pedagogy would allow students to identify scientific information that is relevant and useful rather than focusing on decontextualized facts (Feinstein, 2014).

Although groups such as the National Research Council (NRC, 2002) and the National Science Board (NSB, 2006) have called for improving the assessment practices of student scientific knowledge to include overall student understanding of science concepts as well as subject matter, critics claim that schools still focus almost exclusively on content (Rudolph, 2014). Further criticism points to the lack of educational experiences that allow students to test their own hypotheses, present their findings, and critique each other's findings (Feinstein, 2011).

Student Attitudes Towards Science

Proponents for reconceptualizing science education assert that knowledge alone is not sufficient to make a person scientifically literate, but that one also needs the necessary motivation and beliefs to engage with that knowledge on a daily basis (Fives, Huebner, Birnbaum, & Nicolich, 2014). However, many students do not find their science classes interesting or motivating because of curricula that is overloaded with content that emphasizes isolated facts (Hofstein, Eilks, & Bybee, 2011). Addressing this lack of interest is important because a wealth of research has consistently shown the positive links among student attitudes, motivation, learning, and behavior (Osborne, Simon, & Colling 2003). In fact, research suggests that a deep personal interest in any field of science provides students with the motivation to interact with other scientific fields in the future, and allows students increased confidence in their ability to learn science (Feinstein, Allen, & Jenkins, 2013).

More recently, conceptualizations of science literacy have included student attitudes and beliefs. Self-efficacy, or a student's beliefs about his or her personal capability to succeed, plays a particularly influential role in science education in that research has indicated a significant connection between a student's beliefs about his or her personal competence to succeed in science and academic outcomes (Britner & Pajares, 2001). Conversely, students who lack the

belief they can succeed in science related activities, are more likely to avoid them and give up when faced with obstacles (Britner & Pajares, 2006).

Students' perceptions of their efficacy, or capability to accomplish a particular task has also been linked to the value they place on the task itself. This has led some researchers to assert that efficacy beliefs and task value should be studied concurrently (Zepeda, Richey, Ronevich, & Nokes-Malach, 2015). Furthermore, studies have found that instructional practices can improve motivation and increase task value for science learning (Richey, Ronevich, & Nokes-Malach).

Also predictive of higher levels of achievement and learning are student's epistemological beliefs (Sinatra, Kienhues, & Hoffer, 2014). As with motivation and task value, epistemological beliefs have been shown to be influenced by classroom instruction. Evidence suggests that hands-on classroom experiences that involve students in experimental design, data collection, and analysis may promote more sophisticated epistemological thinking (Solomon et al., 1996). Research has also indicated that constructivist classroom practices may result in epistemologist views that focus on the role of ideas in the process of gaining knowledge, rather than a belief that knowledge acquisition occurs in a piecemeal fashion by making observations, and fails to link the role of ideas to the process (Smith et al., 2000).

Measuring Student Attitudes Toward Science

Advocates for teaching methods that incorporate student motivation and beliefs about science learning have argued for the inclusion of affective elements within assessments (Osborne, Simon, & Colling, 2003). Despite this call for inclusion, existing measures of science literacy do not address student motivation and beliefs (Holbrook & Rannikmae, 2007). Furthermore, few measures of student attitudes toward science have sufficient psychometric data to serve as reliable sources of information (Osborne, Simon, & Colling, 2003). Researchers have

expressed concern with the quality of measures used to assess student attitudes toward science and have noted a lack of instruments that are both valid and reliable (Blaylock et al., 2008).

Rationale and Purpose of the Study

Educators, scholars, and researchers have called for the reconceptualization of what it means to be a scientifically literate person, advocating for assessing scientific learning as a way of thinking that includes student motivation and beliefs toward science. However, the literature points to a lack of measures for assessing students' science literacy within this broader context.

The Science Literacy Assessment (SLA) was developed to assess middle school students' ability to think scientifically and students' motivation and beliefs toward science (Fives, Huebner, Birnbaum, & Nicolich, 2014). Developers of the SLA defined scientific literacy as the "ability to understand scientific processes and to engage meaningfully with scientific information available in daily life" (Fives et al., 2014, p. 550). This definition focuses more on processes that span disciplines, allowing individuals to engage with science in practical and meaningful ways within daily life. The SLA is unlike other measures in that it was designed to measure science literacy as a way of knowing for middle school students. It also fills a gap with the inclusion of an assessment of students' motivation and beliefs about science. The addition of the motivation and beliefs component, allows for the assessment of science literacy within a broader context. It also addresses a limitation of most existing measures of science literacy, which is the exclusion of assessment of motivation and beliefs, despite research indicating student attitudes and beliefs play an important role in engaging with science (Fives, et al., 2014).

The SLA has two parts: the first assesses five components of *demonstrated* science literacy (SLA-D); the second is a motivation and beliefs scale (SLA-MB) that measures students' self-efficacy, subjective task value, and personal epistemology for science. There are three

versions of the SLA-D; two contain 26 multiple-choice items, and the third is a shortened 19-item version. The SLA-MB is composed of 25 Likert items that comprise the three subscales; self-efficacy for science literacy, the value of science, and personal epistemology.

While the developers of the SLA provided validity evidence for the 26-item version related to test content, response process, and internal structure, this study aimed to confirm the validation argument by replicating evidence based on internal structure for the 19-item version. The study also aimed to expand the validity evidence to include convergent validity, and predictive validly for the 19-item version. And finally, validity evidence will be further expanded by examining how the measure performs when given to high school students.

Research Questions

The following research questions guided the current validity study for the Science literacy Assessment:

1. To what degree is the SLA (SLA-MB; SLA-D 19 item version) a valid measure of middle school students' general scientific knowledge, motivation, and beliefs related to science?
2. To what degree is the SLA (SLA-MB; SLA-D 19 item version) a valid measure of high school students' general scientific knowledge, motivation, and beliefs related to science epistemology?
3. To what extent do scores on the SLA-MB correlate with scores on The Modified Attitudes Toward Science Inventory (mATSI) and the STEM Career Interest Survey (STEM-CIS)?
4. To what degree do scores on the SLA-D predict scores on a student general science knowledge test?

Definition of Terms

Accountability Policies: Policies that mandate tests with results made publicly available (Anderson, 2012).

Attitude: A general and enduring positive or negative feeling about a person, object, or issue (Petty & Cacioppo, 1996).

High Stakes Testing: The existence of possible sanctions for schools if they do not meet set benchmarks indicating success (Anderson, 2012).

Nature of Science: The epistemology of science, science as a way of knowing, or the values and beliefs inherent to scientific knowledge and its development (Lederman, 1992).

Scientific Self-efficacy: A person's beliefs about his or her personal competence to succeed in science related tasks (Zusho & Pintrich, 2003).

Standards Based Education: Systems of instruction, assessment, grading, and academic reporting that are based on students demonstrating mastery of the knowledge and skills they are expected to learn as they progress through their education (Great Schools Partnership, 2004).

Task value: A measure of worth placed upon completing a given task. Task value can be broken into four categories; attainment value refers to the importance students' place on doing well, student's interest in the task is referred to as intrinsic value, utility value relates to the task's level of importance for the students' current and future goals, and cost can be defined as negative consequences associated with engaging in the task (Eccles & Wigfield, 2002; Pintrich and Schunk, 2002)

Personal Epistemology: The ways in which a person thinks about how knowledge is gained and the sources of that knowledge (Sinatra, Kienhues, & Hoffer, 2014)

Chapter Two

Review of the Literature

This chapter begins with a summary of common justifications for advancing the public's understanding of science. Next is an overview of how the concept of science literacy has evolved in the literature and has been influenced by goals for promoting science literacy. A summary is then provided of science education reform movements in the United States to provide a context for the changing nature of science literacy and how shifts in goals and concepts have influenced policy. The review then examines ways in which evolving policy has driven the ways in which the construct of science literacy has been measured. The chapter then concludes with a description of extant literature describing the link between student attitudes toward science and achievement as correlated with science literacy constructs.

Method for Review of the Literature

The search strategy for this review of the literature was conducted using four procedures: (a) electronic search of literature databases, (b) manual search of published literature, (c) electronic search of dissertations, and (d) electronic search of reports and documents created by organizations associated with K-12 education in the United States such as the National Science Teachers Association, and the American Association for the Advancement of Science.

An electronic search of Dissertation Abstracts Complete, Academic Search Complete, Education Research Complete, PsycINFO, JSTOR, and ERIC was conducted without limitations placed on date of publication in order to capture literature which would provide a complete historical perspective of science literacy in the United States. Combinations of relevant search

terms were utilized within each database. Specific terms included: *science literacy, scientific literacy, science education, science educational reform, nature of science, student attitudes toward science, student motivation, literacy, science literacy measurement, and science literacy instrument*. Similar combinations of key words were then use to search the ProQuest dissertation database. Again, no restrictions were placed on publication dates. Reference lists were examined from literature resulting from these searches, yielding additional sources for consideration. This additional literature included scholarly work which appeared in peer-reviewed journals or conference proceedings, textbooks, accepted dissertations, and reports from government-funded educational organizations. Titles and abstracts were reviewed for each relevant publication and copies of primary sources were obtained for each source found to be applicable. These sources were then vetted using the *Standards for Reporting on Empirical Social Science Research in AERA Publications* (AERA, 2006). Inclusion criteria based on AERA standards include, but are not limited to, transparent reporting of activities that guided empirical research as well as the provision of adequate evidence to justify results and conclusions.

Justifications for Promoting Science Literacy

A number of arguments exist for promoting the public's understanding of science. These justifications address topics including economics, scientific advancements, military superiority, ideology, culture, intellect, aesthetics, and ethics (Shortland, 1988). Laugksch (2000) groups these arguments into two categories. The first represents a micro view and relates to how the lives of individuals are enhanced through a greater understanding of science. The second category, a macro view, relates to the benefits science literacy brings to the nation, science, or society.

Benefits of Increased Science literacy at the Micro Level

Arguing for the necessity of science literacy for individuals, Thomas & Durant (1987) contend that improved understanding of science and technology is beneficial to anyone living in a science- and technology-dominated society. In fact, the need to understand scientific concepts and to make sense of sometimes overwhelmingly vast amounts of information from diverse sources is paramount in modern cultures. Along with the increasing volumes of information comes the challenge of discerning accurate and relevant facts in order to make decisions relevant to life choices. Sinatra, Kienhues, and Hofer (2014) believe citizens with greater science literacy are more able to effectively and confidently negotiate their way through the society in which they live. They provided examples to support this belief, including; an individual may choose to spend more for a car that is fuel efficient, vote based on issues such as climate change and requirements of labeling genetically modified organisms, and decide whether it is beneficial or harmful to give their small child an electronic device.

Other arguments for promoting science literacy relate to the intellectual, aesthetic, and moral benefits to individuals (Laugksch, 2000). Shortland (1988) suggests that the promotion of science literacy encourages a more intellectual culture, expressing that "science is the distinctively creative activity of the modern mind" (p. 310). The author asserted that science is as central to a truly cultivated mind as literature, music, and performing arts. This argument proposes that we should support science literacy for the same reasons that we preserve beautiful buildings and paintings (Laugksch, 2000).

Being competitive for employment is also a benefit to individual citizens. As global economies become more knowledge-based, the quality of a nations' workforce is increasingly viewed as one of the most important economic assets of modern science and technology-based

societies (Brooks, 1991). Scientifically literate individuals may therefore be in a better position to secure new job opportunities and be more able to utilize technology in the workplace (Thomas & Durant, 1987). This view is supported by the latest National Science Board Indicators (2016), a report that provides comprehensive federal data on a wide range of measurements related to science and engendering. Between 1960 and 2013, the number of workers in science and engineering occupations grew an average of 3% per year, compared to a 2% annual growth rate for the total workforce. In 2013, approximately 5.7 million college graduates in the United States entered employment in science and engineering occupations. Additionally, unemployment rates for college-educated individuals in science and engineering occupations tend to be lower than rates for all college graduates and significantly lower than those for the overall work force: In February 2013, the unemployment rate for scientists and engineers was around 3.8%, which was less than 4.3% of all college-educated individuals in the labor force which were unemployed, about half the 8% official unemployment rate for the entire U.S. labor force (NSB, 2016). These indicators point to the reason policymakers increasingly emphasize the important role of science and technology in a country's economic growth and competitiveness.

Benefits of Increased Science literacy at the Macro Level

Shortland (1988) suggested that a high level of science literacy within the general population produces greater support for science itself. This occurs in part, because an awareness of the work done by scientists attracts new recruits into the field. Shortland argued that unless the general public values what scientists are attempting to achieve, science is unlikely to be financially supported. In other words, if the general public understands the objectives, processes, and capabilities of science, they are more likely to appreciate the usefulness of scientific endeavors and support research efforts.

An argument closely related to public support for science is the public's right to participate in science policymaking and the expenditure of public funds (Thomas & Durant, 1987). Prewitt (1983) supported this argument and contends that scientifically literate citizens help to strengthen the democratic process through meaningful involvement with political processes, public policymaking, and social change. The desire for public support for military spending is also linked to science literacy. Anelli (2011) contends that the impact of the first Sputnik launch on the educational system in the United States cannot be overemphasized, stating "Stunned by the Soviet Union's achievement, which conceivably could be translated into ballistic missiles carrying nuclear weapons from Europe to the U.S., our government responded by pouring billions of dollars into science education" (p. 236).

Perhaps the most frequently expressed rationale for improving science education stems from the view that a scientifically literate populace as necessary for the economic well-being and security of the nation. In an ever increasingly interconnected world, economic competitiveness depends on a nation's capacity to successfully compete in international markets. This requires a strong national research and development program as well as a steady supply of scientists, engineers, and technically trained workers (Laugksch, 2000). Given that workers with scientific skills are vital to a nation's economic competitiveness, only nations whose citizens possess an adequate level of science literacy will be able to sustain the need for an ongoing need to supply such workers. Therefore, when studies critique the quality of public education, government leaders tend to perceive problems with science education as a threat to the future economic strength of the United States and its position as a global leader (Anderson, 2012; National Science Board, 2006). Pervasive anxiety about the United States' ability to produce workers with scientific skills has caused many to express concern that this country is losing ground

within a global economy. In fact, among countries with large numbers of researchers, growth since 2000 has been most rapid in China and South Korea. The United States and the European Union experienced steady growth but at lower rates than China or South Korea (National Science Board, 2016). Additionally, global public investment in research, development, and alternative energy totaled an estimated $12.7 billion in 2013. The European Union invested the most, spending $4.4 billion, followed by the United States ($3.5 billion), Japan ($2.6 billion), and Canada ($0.8 billion) (National Science Board, 2016). At 39%, the United States leads the world in the percentage of its Gross Domestic Product that comes from knowledge-intensive service industries and high-technology manufacturing. In high technology manufacturing, the United States retains a slim lead as the largest global provider (29%) over China (27%), whose global share rose steeply since the turn of the century (National Science Board, 2016).

Conceptualizations of Science Literacy

Education, being a social activity, is influenced by ideology. Achieving a scientifically literate populace is widely acknowledged as an important goal and is often claimed to be the desired outcome of science education; however there is no consensus on exactly what that means (DeBoer, 2000). In fact, the U.S. has undergone continuous science educational reform efforts since the end of World War II (Baybee, 1997). Although the term ‘science literacy’ is central to discussions of policy, research studies, and curriculum reform efforts, the construct has remained somewhat ambiguous and has lacked universal definition since its introduction in the 1950s (DeBoer, 2000; Liu, 2009; Roberts, 2007). The collective body of literature defining science literacy has been described as unwieldy, lacking focus, and producing multiple ways to conceptualize the construct (Roberts, 2007; Wenning, 2006).

Asserting that most definitions of science literacy have focused on identifying what is valuable for students over the course of their lifetime, regardless of their career goals, Roberts (2007) identified several approaches used by authors to define the concept. Four of Roberts' approaches are utilized to present the definitional literature within this study; the first consists of attempts to synthesize a definition by analyzing how the term science literacy was used within education literature from around 1960 through 1980. A second approach concentrates on types or levels of science literacy based on the assumed needs of learners. A third approach focuses on conceptualizations resulting from the input of scientists, and the fourth approach focuses on the contexts in which science is considered to be important in the everyday lives of students.

Analyzing Educational Literature to Synthesize a Definition of Science Literacy

An example of Roberts' first approach, attempts to synthesize a definition by analyzing how the term was used in early literature, is the work of Pella, O'Hearn, & Gale (1966). The authors conducted a comprehensive literature search in order to examine how the term was used within scientific literature. The authors searched for titles that included (a) science literacy, (b) science and/or technology and the citizen, (c) relationship of science and technology, (d) relationships or interrelationships of technology and/or science and society and social problems, (e) science and/or technology and culture, (f) relationship between nonscientists and scientists, (g) science and the public domain, and (h) the technological and/or scientific revolution. The authors identified 100 papers for analysis and each "was carefully studied in terms of the following questions: (a) What does the author mean by science literacy, science for the citizen, science for general education, etc.? (b) Does the author indicate referents to science literacy? If so, what are the referents? (c) Does the author discuss scientific and/or technological needs of citizens? If so, what are these indicated needs." (p.199).

They reported the six most frequent referents to science literacy as relating to:

1. interrelations between science and society (67 references)
2. ethics of science (58 references)
3. nature of science (51 references)
4. conceptual knowledge (26 references)
5. science and technology (21 references)
6. science in the humanities (21 references)

Based on their examination, the authors concluded that a scientifically literate individual was characterized as one with an understanding of the basic concepts in science, the nature of science, ethics that guide the work of scientists, interrelationships of science and society, differences between science and technology, and interrelationships between science and the humanities. They also found that the primary purposes of science literacy were to adequately prepare students who were aiming to obtain careers in science, and to provide all students with the appropriate scientific general education required for effective citizenship.

Gable (1976) also utilized the method of examining educational literature for his dissertation research in order to develop a theoretical definition of science literacy. He reviewed literature published between 1885 and 1976 in order to identify statements describing science literacy. The result of Gable's work was the development a theoretical model that conceptualizes science literacy as a two-dimensional matrix containing 72 cells. One dimension expanded the categories proposed by Pella et al. and included components such as intellectual process, science concepts, nature of science, and relationships between science and society. Gable found the most referents, 67, within the category of science and society, followed by the ethics of science, having 59. Gable noted that the literature was couched in terms of behaviors

that could be expected of a scientifically literate person. Therefore, the second dimension of the theoretical model incorporated cognitive and affective categories of Bloom's (1956) Taxonomy, including the six major categories of cognitive objectives (knowing, understanding, applying, analyzing, synthesizing, and evaluating) and five categories of affective objectives (receiving, responding, valuing, organizing, characterizing).

DeBoer (2000) used a slightly different approach to using literature to examine science literacy. DeBoer wanted to understand how historical events influenced science education policy statements over time (Roberts, 2007). DeBoer (2000) conducted a historical analysis of standards-based reform efforts within the United States. He identified nine goals for teaching science that represent "a wide range of meanings of science literacy" (p.591). These goals included: (a) teaching and learning about science as a cultural component of the modern world, (b) preparation for the world of work, (c) teaching and learning about science that has direct application to everyday living, (d) teaching students to be informed citizens, (e) learning about science as a specific way of examining the natural world, (f) understanding reports and discussions of science that appear in the popular media, (g) learning about science for its aesthetic appeal, (h) preparing citizens who value science, and (i) understanding the nature and importance of technology and the relationship between technology and science. DeBoer proposed that reviewing educational history reveals that science literacy is a general concept that has a wide variety of meanings. The primary conclusion that can be drawn from DeBoer's study is that the term science literacy has usually implied a broad and functional understanding of science for general education purposes and not necessarily for preparation for specific scientific and technical careers. According to DeBoer (2000), science literacy is comprised of knowledge the public should have about science in order to live in ways that respect the natural world.

Types and Levels of Science literacy Based on the Assumed Needs of Learners

Roberts (2007) second approach to defining science literacy focuses on differences instead of commonalities for the purpose of creating categories that differentiate types of science literacy according to what learners should be able to do. Shen (1975a) proposed six features of science literacy. They include an individual's ability to understand (a) basic science concepts, (b) the nature of science, (c) ethics guiding scientists' work, (d) interrelationships between science and society, (e) interrelationships between science and humanities, and (f) relationships and differences between science and technology. Based on these six features, Shen (1975a) proposed three distinct types of science literacy that differ in the content and format of objectives, the target audience, content, and modes of delivery. The first, practical science literacy, means possessing scientific knowledge than can be used to help solve everyday problems and improve standards of living. The second type, civic science literacy, is the ability of a citizen to become more aware of science, and science-related issues, in order to participate in the democratic process. The third level, cultural science literacy, is motivated by an individual's desire to gain knowledge and appreciation of science as a major human achievement. Shen (1975b) argued that practical literacy is the most urgently needed and most frequently neglected type of science literacy, citing a global lack of knowledge of health, nutrition, and modern agriculture that could ease human suffering.

While Shen's three categories represent differing types of science literacy, they were not arranged in any form of hierarchy. Shamos (1995) extended Shen's work, proposing that different amounts of science literacy are necessary for achieving Shen's three types of science literacy, thus converting them to levels in a hierarchy. Shamos argued that assuming an individual is either scientifically literate or illiterate is an oversimplification. Instead he

categorizes science literacy into three levels, each building upon the other. These levels include (a) cultural science literacy; defined as grasping certain background information required for basic communication, including science-related terms, (b) functional science literacy: where one not only knows science terms, but also is able to converse, read, and write coherently using these terms in non-technical but meaningful ways, and (c) true science literacy: which demonstrates an overall understanding of science and the major theories that form the foundation of science. This level also requires an appreciation of scientific investigation and the importance of questioning, analytical reasoning, logical thought processes, and relying upon objective evidence.

Similarly, Bybee (1997) created a framework that presents science literacy as a continuum in which individuals develop greater and more sophisticated understanding of science and technology across four domains. At the first level, nominal literacy, an individual is able to associate names with general areas of science and technology without necessarily having an accurate understanding of the concepts. The second level, functional literacy, refers to a person's ability to read and write passages with simple scientific vocabulary. To obtain the third level, conceptual and procedural literacy, one must demonstrate an understanding of science as a discipline as well as the procedures for developing new knowledge. To achieve Bybee's fourth level of science literacy, multidimensional literacy, requires an understanding not only the structure of science and technology, but also the nature of science and technology and their relationships with society.

Conceptualizations of Science literacy Originating from Scientists

Roberts (2007) contends that asking scientists to recommend the essential subject matter for school science has often been a part of science education. For example, Project 2061 was created by the American Association for the Advancement of Science (AAAS) in 1985 when

Haley's comet last passed close to earth. The project was named for the year of the next return of the comet, representing the realization "that children who would live to see the return of the Comet in 2061 would soon be starting their school years" (AAAS, 1992, p. 5). Project 2061 developed from the work and recommendations of five panels appointed by AAAS, composed primarily of scientists who were charged with developing recommendations for educational reform in five areas: biological and health sciences; mathematics; physical and information sciences and engineering; social and behavioral sciences; and technology (Eisenhart, 1996). The Project 2061 conception of science literacy was constructed initially from five reports developed between 1985-1989 under the guidance of the National Council on Science and Technology Education. The reports contained the "understandings and habits of mind" the group considered essential for all citizens in a scientifically literate society (AAAS, 1989, p.3). Following the publication of the reports, the council solicited broad consultation and review. The process involved hundreds of individuals and resulted in the sixth report of the collection, Science for All Americans (SFAA), which was unanimously approved by the AAAS Board of Directors.   The AAAS definition of science literacy encompasses mathematics and technology as well as the natural and social sciences. According to AAAS, a scientifically literate person is one who is aware that science, mathematics, and technology are interdependent human enterprises with strengths and limitations, who understands key concepts and principles of science, who is familiar with the natural world and recognizes both its diversity and unity, and who uses scientific knowledge and scientific ways of thinking for individual and social purposes.

Defining Science literacy Within the Context of Everyday Life

Feinstein (2011) argued that a truly useful conceptualization of science literacy must relate to the use of science within daily life, sometimes referred to as public engagement with

science. By this he is supporting the notion that science education can help people solve personally meaningful problems, inform their behavior, and influence their political decisions. This conceptualization of science literacy appeared as early as 1947 in a report, *Science Education in American Schools*, produced by the National Society for the study of Education (NSSE). Within this report, the authors asserted that science concepts and principles must be taught so they will be "functional" (p.26), stating:

> The critical element in functional concepts and principles, as in functional information, is understanding. Understanding is not quickly achieved. It rarely results in any useful amount from a single experience or from exact duplicates of that experience how often it is repeated… For the kinds of concepts and principles which are properly science objectives there must be many and varied experiences in which the same idea, large or small, occurs in differing situations. Moreover, for the most fruitful learning, these experiences must be arranged and graded with respect to complexity and difficulty, so that the pupil may be guided to organize his meanings at higher and higher levels. Meaning learning is spiral. Each experience adds a new loop in the spiral of meaning (p.27).

Focusing more on the individual's role within a larger society, the science education community became focused on the role of scientific knowledge on society in the 1960s (DeBoer, 2000). During this period, scientists such as Polykarp Kusch (1960), a physicist, called for a broader conception of science and drew a connection between scientific knowledge and good citizenship. This relationship between science and society was promoted as a goal of the science curriculum and gained additional momentum when the National Science Teachers Association

(NSTA) identified it as the most important goal of science education in its position statement *School Science Education for the 1970s*.

Expanding upon the concept of the roles individuals play within society, Layton, Davey, and Jenkins (1986) offer a more situation specific concept of science literacy. They argue that the scientific knowledge people use is contextualized according to what information the situation requires. These authors address the meaning and social uses that science has for the general public, introducing the term *science for specific social purposes* to stress the idea that the context or situation of a socio-scientific issue has a strong influence on the knowledge people use to address it. Also advocating for a contextually relevant view of science literacy, Eisenhart, Finkel, and Marion (1996) encourage what they term “socially responsible science use” (p.283), a concept that involves learning science within real life situations or scenarios. These authors maintain that teaching students key concepts and scientific methods of inquiry do not necessarily lead to the socially responsible use of science, instead they proposed the use of science in socially responsible ways would involve: (a) understanding how science related actions impact individuals who engage in them; (b) understanding the impact of scientific decisions on others, the environment, and the future; (c) understanding relevant scientific content and methods; and (d) understanding the advantages and limitations of a scientific approach.

Building on the perspective of Eisenhart et al., Roth and Lee (2002, 2004) proposed rethinking science literacy as more of a social practice. They promoted three ideas. First, not everyone needs to learn the same set of basic concepts because society is built on a division of labor, therefore it is more important to allow science literacy to develop as a collective, social phenomenon. Second, scientific knowledge should not be overemphasized in democratic decision making, but should be one of many factors in an approach that considers other factors

such as politics, economics, and ethics. Third, presenting science education in ways that allow students to participate in community life encourages the potential for lifelong participation in and learning of science-related issues. Roth and Lee presented these views following a three year multi-site ethnographical study which was conducted both in schools and in the community, focused on a variety of community events that arose from concerns about addressing water quality problems. Over a two-year period, Roth and Lee co-taught with science teachers in three seventh grade classes in which students designed and conducted their own projects relating to a local creek with the intent of reporting their findings at an open-house event. The students were encouraged to choose the data collection and reporting tools that best fit their interests. Other students within the community conducted research in the watershed as part of their involvement in science fair competitions. Community members were an integral part of the process. For example, every other week the classes spent an afternoon in and around the creek. Community members assisted with activities such as driving children to different sites along the creek, teaching lessons, and assisting students with analyzing data on organisms obtained from the creek. Students from classes that had already completed their projects supported their peers who were just beginning by assisting with fieldwork and data analysis. Through this process, the students exchanged and produced knowledge with other members of the community. Roth and Lee found that the activities of the students mirrored the activities of other community members in addressing the water-related problems. For example, both groups set goals, identified tools, assigned division of labor, and established rules for interaction. Similarly, findings indicated that collectively, more advanced forms of science literacy were produced than any one individual could produce. Children generally participated in the activities with similar motives as the adults, and they participated in various forms of conversation with adult community members.

Based on this qualitative research, the authors put forth a concept of science literacy that emphasizes a collective, rather than individual, quality. They asserted that children who participate in activities that contribute to the knowledge within their community can continue this participation throughout their lifetime.

More recently, Feinstein (2011) expanded the notion of basing educational strategies on an understanding of the role science plays in everyday life by advocating for a connection between science education research and the broader field of public engagement with science. He believes that if we want students to be attentive to scientific issues attached to major personal and political decisions, we should learn how laypeople become involved in science. Feinstein asserted that people integrate scientific ideas with their personal experiences in order to draw conclusions and make decisions. He further argued that people do not remove themselves from their own personal context when engaging in science in order to place themselves into the role of a scientist. He refers to scientifically literate laypeople as *competent outsiders*, indicating that these nonscientists are able to recognize when science has relevance to their lives and to interact with sources of scientific information in order to help them achieve a goal.

Science Education Reform Efforts in the United States

Embedded within social context, education policy is shaped by the shifting of public perspectives and goals across time. In presenting a historical overview of large scale reform efforts in U.S. science education, Kahle (2013) identifies three waves of systemic reforms. The first wave covers the 1950s through the 1970s and was sparked by the launch of Sputnik in 1957. The second wave was signaled by the release of *A Nation at Risk*, a 1983 report issued by the National Commission on Excellence in Education, which highlighted the falling scores and academic underachievement of American students. The third wave was marked by the

establishment of the Statewide Systemic Imitative (SSI) program in 1990 by the National Science Foundation which acknowledged for the first time a need to address whole state educational systems when working toward change.

Wave I: Science Education Reform following World War II

At the conclusion of World War II, Americans began refocusing their attention on domestic issues such as education. Several areas within the U.S. educational system were in need of significant improvement. For example, repairs to schools and classrooms had been delayed during the Depression and the war. Also, "baby boomers" were expected to overcrowd existing facilities, and America had entered an age driven by science and technology in which citizens would need higher levels of education to sustain a technologically oriented society and compete economically (Baybee, 1997). In a 1950 report by the Harvard Committee, *General Education in a Free Society*, the committee recommended that social studies, science, and mathematics be included in the secondary curriculum. The report emphasized general education as "that part of a student's whole education which looks first of all to his life as a responsible human being and citizen" (p.51). The report also supported ability grouping, suggesting each of the disciplines be "courses of different difficulty" (p. 100) with the only requirement for enrollment in one of the courses being ability.

The first wave of science education reform aimed to address two additional areas of weakness, outdated texts and the quality of science teachers. From 1954 to 1974, the National Science Foundation funded Teacher Training Institutes which were attended by over 40,000 teachers each year. The Institutes brought teachers up to date with current scientific developments, allowed teachers to conduct experimental work in science and encouraged them to replicate these experiments with their students, and provided a network of peers for

professional support (Kahle, 2013). In order to address significantly outdated textbooks which lacked information on current scientific advances, a group of physicists in Cambridge formed the Physical Science Study Committee (PSSC) in 1954 and began developing new curriculum. By the late 1950s curricula had been developed in physics, chemistry, biology, and earth sciences. However, most of this curricula development was led by scientists and scholars with a bias toward their own disciplines and whose main focus was on a need to produce more scientists and engineers. Also, the curricula developed were mostly discipline-focused and designed for above-average students (Harvard Committee, 1950; Kahle, 2013).

During the 1960s and 1970s, educational reform mirrored the tremendous social change occurring in the United States. Within the Civil Rights and Women's movements, groups that had previously been subjugated began to demand equity and opportunity for all citizens (Airasian, 1987). In 1965, President Lyndon B. Johnson signed into law the landmark Elementary and Secondary Education Act (ESEA) which established, for the first time, federal funding for public education combined with a federal policy that emphasized equal access to education and supported educational opportunities for students from high poverty communities (Forte, 2010).

Accompanying the push for equality was a growing belief that human behavior is pliable, and that social problems such as poverty, discrimination, and educational disadvantage, could and should, be changed (Airasian, 1987). The resulting major reform initiatives during this period required that educational systems served a variety of groups that previously were either underserved or not served at all. Furthermore, beliefs that people are most responsive to environmental stimuli when they are young produced reform efforts focused on preschool and early elementary school programs (Bloom, Davis, & Hess, 1965).

Wave II: Science Education Reform Since the 1980s

The second wave of U.S. science education reform was marked by a 1983 report by the National Commission on Excellence in Education (1983) issued a report, *A Nation at Risk: The Imperative for Educational Reform*. The report argued that academic standards had fallen in the U.S. and pointed to low test scores of American youth, especially in math and science. This poor academic performance was said to be the cause of our declining economic position in the world and the recommended solution was to create a more rigorous academic curriculum for all students built around the basic academic subjects of English, mathematics, science, and social studies, as well as computer science and foreign languages. The more rigorous curriculum would be accompanied by higher standards for all students and new methods of assessment and accountability (DeBoer, 2000). A parallel paradigm shift in federal education policy was echoed within the 1994 reauthorization of the ESEA, which required states to (a) establish common, statewide standards for all students in reading and mathematics in the 3-5, 6-8, and high school grade ranges; (b) implement statewide assessments aligned to these standards in at least three grades each for reading and mathematics; and (c) implement a statewide accountability system for evaluating school-level performance (Forte, 2010). The subsequent 2001 reauthorization of ESEA, No Child Left Behind (NCLB) extended these requirements to include standards and assessments in at least three grades for science. Forte (2010) describes the logic behind this approach as being: If (a) performance standards clearly define what students should know and be able to do as well as the cognitive level to which students should be able to demonstrate this knowledge within each grade level and content area, and (b) assessments are aligned to these expectations, then (c) assessment scores can be used to inform accountability decisions meant to improve school quality and student achievement.

The heightened sense of urgency created by the Nation At Risk report coupled with federal standards-based reform efforts resulted in the development of three national-level proposals for science education reform, including the National Science Teachers Association (NSTA) Project on Scope, Sequence, and Coordination of Secondary School Science (Aldridge, 1992; NSTA, 1992, 1995); the American Association for the Advancement of Science's (AAAS) Project 2061 (AAAS, 1993; Rutherford and Ahlgren, 1990); and the National Research Council's (NRC) National Science Education Standards (1994). This group of major educational reforms had as their goal the development of an increased degree of science literacy among school children, and ultimately, the general public (Wenning, 2006).

NSTA Project on Scope, Sequence, and Coordination of Secondary School Science. NSTA's Project on Scope, Sequence, and Coordination of Secondary School Science (SS&C) was initiated in 1998 and reflects the interests of NSTA's membership, science teachers, science education faculty, and educational administrators, rather than scientists. The purpose of SS&C is to increase levels of science literacy by reforming the way science education is organized and science education is taught (NSTA, 1992). SS&C focuses on recent developments in learning theory and science education research, including provisions for hands-on experience, sequencing over time at successively higher levels of abstraction, and taking account of student pre-conceptions (Aldridge, 1992). At its heart, SS&C advocates "a reform project that places learning over several years and moves from concrete experiences to abstraction. Using a spiral approach, the same concepts, principles, laws, and theories are studied at successively higher levels of abstraction, thus helping students to construct their own knowledge" (Aldridge, 1992, p.17).

AAAS Project 2061. First published in 1989 by the American Association for the Advancement of Science (AAAS), Project 2061's *Science For All Americans* called for reform efforts toward standards based education, which includes systems of instruction, assessment, grading and academic reporting that are based on students demonstrating mastery of the knowledge and skills they are expected to learn as they progress through their education (Great Schools Partnership, 2004). The purpose of the report was to clarify the goals of science education so that all students could attain science literacy (DeBoer, 2000). The premise for reform was based on the perception that the U.S. had not responded as quickly as other countries in preparing its youth for a world in which science and technology play such a large part. This call for reform would require a consensus on what students needed to know to be scientifically literate (Wenning, 2006) and was a reaction to the assertions made by the Nation at Risk report, which claimed that student achievement in the United States was falling behind that of other nations because of inadequacies within the educational system (Gardner 1983).

NRC National Science Education Standards. Following the call for standards based education put forth by AAAS; the National Science Teachers Association proposed in 1991 that the National Academy of Sciences and its research arm, the National Research Council (NRC), construct a set of national standards for science education. As a result, the NRC published a draft of its *National Science Education Standards* in 1994 (Eisenhart, 1996). The NRC standards content reflected unified science concepts and processes, science as inquiry, physical science, life science, earth and space sciences, science and technology, science in personal and social perspectives, and the history and nature of science. According to the NRC, scientific proficiency consists of four strands: (a) knowing, using and interpreting scientific explanations of the natural world; (b) generating and evaluating scientific evidence and explanations; (c) understanding the

nature and development of scientific knowledge; and (d) participating productively in scientific practices and conversations (NRC, 1996). The objective of the National Standards is for all students to achieve science literacy by mastering a common set of content standards (DeBoer, 2000). There are five main assumptions underlying the identification of the content standards: (a) "Everyone needs to use scientific information to make choices that arise every day." (b) "Everyone needs to be able to engage intelligently in public discourse and debate about important issues that involve science and technology." (c) "Everyone deserves to share in the excitement and personal fulfillment that can come from understanding and learning about the natural world." (d) "More and more jobs demand advanced skills, requiring that people be able to learn, reason, think creatively, make decisions, and solve problems. An understanding of science and the process of science contributes in an essential way to these skills." (e) "To keep pace in global markets, the United States needs to have an equally capable citizenry" (National Research Council, 1996, pp. 1-2).

Implementation of the National Standards was not without criticism. The standards were developed by a wide range of individuals representing many constituencies (Collins, 1998), as a result, the definition of science literacy has been criticized as broad and inclusive of virtually all objectives of science education that have ever been identified (DeBoer, 2000). Eisenhart, Finkel, and Marion (1996) characterized the implementation of the Standards (1994) as focusing "narrowly on key content: specifying what facts, concepts, and forms of inquiry should be learned and how they should be taught and evaluated" (p.266). The authors contend that there is an assumption inherent in the standards "that producing citizens who can use science responsibly and including more people in science will naturally follow from teaching a clearly defined set of scientific principles and giving students opportunities to experience 'real' science" (p. 268).

They also maintain that teaching students key concepts and scientific methods of inquiry does not necessarily lead to the socially responsible use of science or to an increase in the number of citizens who participate in discussions of scientific issues.

Some of these criticisms were addressed in 2012 when the National Academy of Sciences developed a structure to standardize K-12 science education. Entitled *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*, the publication encourages science educators to focus on a limited number of disciplinary core ideas and crosscutting concepts, be designed so that students continually build on and revise their knowledge and abilities over multiple years, and support the integration of such knowledge and abilities with the practices needed to engage in scientific inquiry and engineering design (NRC, 2012). The following year new standards for science education were issued, updating the national standards released in 1996. Developed by 26 state governments and national organizations of scientists and science teachers, the guidelines, called the Next Generation Science Standards (NGSS), are intended to standardize teaching among states, and to raise the number of high school graduates who choose science and technical majors in college (NGSS, 2013).

The development of the NGSS reflects a pattern of standards-based education that has become the focus of education reform in the United States. The National Research Council's *Framework for K-12 Science Education* and the NGSS move teaching away from covering a large amount of isolated facts in order to focus on a smaller number of disciplinary core ideas and crosscutting concepts that are used to explain phenomena and solve problems by engaging students in the process of science (Krajcik, Codere, Dahsah, Beyer, & Mun, 2014). The NGSS raise the performance expectations for students and promote inquiry-based teaching methods,

defining inquiry in science as a process that requires a wide range of cognitive, social, and physical activities (NGSS, 2013). The implementation of the NGSS requires a fundamental shift from previous National Science Education Standards (NRC, 1996, 2001) in two ways: (a) significant increase in the level of higher-order thinking skills that students are expected to master and (b) more incorporation of real-world interaction with science content (Marshall & Alston, 2014). The new standards also require students to engage with science through the processes of modeling, analyzing, and designing. Within this framework, practices and content are combined within core ideas, or crosscutting concepts, that span disciplinary boundaries and unifying the study of science (NGSS, 2013). The standards incorporate critical thinking and real world relevance whereas previous state and national science standards were typically designed almost as a checklist of information to be presented. In contrast, the new performance expectations provide objectives that may require days to master, require students to explore and investigate, and provide evidence-based claims about their observations. Within the previous National Science Education Standards, inquiry was separated from the content standards, using instructional methods which often separated the process of science from the content of science. For example, many teachers taught the scientific method as a separate unit, completely removed from the scientific content to be studied. This was problematic in two ways, first, it presented science inquiry as occurring in a single linear sequence; and second, real world and context specific learning experiences were often not used. The NGSS help to address these problems by providing performance expectations that integrate specific practices with core science concepts (National Research Council, 1996).

Test-based Accountability and Science Education

Current accountability policies require the use of standardized tests to measure student's attainment of state or federal goals and mandate that results be made publicly available. These policies, which aim to help educators consider the needs of all students, have been expanding since the mid-1960s (Anderson, 2012). The use of quantitative measures to evaluate educational quality began with ESEA in 1965. President Nixon further promoted the notion of utilizing quantitative measures to assess educational quality when he addressed the Congress in 1970, emphasizing the need for "measurable" standards (Sirotnick, 2004, p. 150). By 1973 accountability legislation had passed in 27 states (DeNovellis & Lewis, 1974) and the practice of holding schools accountable for student learning through the use of educational testing continued to expand through the 1990s (Anderson, 2012). The reauthorization of ESEA in 1988 required yearly examination of educational progress through student testing in schools and districts receiving Title I funding, and in 1994 ESEA linked content standards and student testing to accountability (Penfield & Lee, 2009).

No Child Left Behind. The No Child Left Behind Act (NCLB), signed into law January 2002, required states to report to the federal government the percentage of students who were proficient in reading and mathematics (and subsequently science), disaggregate those results by subgroups (minority status, poverty, limited English proficiency, and students with disabilities), and make 'adequate yearly progress' toward proficiency for all students (Brookhart, 2013). NCLB required states receiving Title I funds to develop academic content standards in mathematics and science by academic year 2005. These standards were to identify the knowledge and skills students were expected to master and beginning in academic year 2007 schools were required to assess student science performance once in elementary school, once in

middle school, and once in high school. Schools were then held accountable for ensuring that students met the set standards. Schools that failed to make adequate yearly progress faced a range of sanctions that varied from reductions in funding to the dismissal of all teaching and administrative staff (Supovitz, 2009). This reform attempted to reduce educational disparities by drilling accountability down to the level of individual classrooms where teachers were to be held accountable for student test performance. (Sutherland, Settlage, & Brickhouse, 2014).

Every Student Succeeds Act. The Every Student Succeeds Act (ESSA, 2015), passed into law in December 2015, is the most recent reauthorization of the 1965 Elementary and Secondary Education Act. Although ESSA replaces its predecessor, No Child Left Behind, the new law retains the current schedule of federally mandated statewide assessments: requiring reading/language arts is to be assessed yearly in grades three through eight, and once in grades nine through 12; and science being assessed at least once in grades three through five, once in grades six through nine, and once in grades 10 through 12. Although the new law will not become effective until the 2018 school year, it allows a for a more robust system of assessing student learning in requiring states to include measures that assess higher order thinking skills as well as permitting states to administer either a single summative assessment or multiple interim assessments throughout the year that result in a single summative assessment.

Science education policy and science instruction. In 2004, NCLB greatly increased the focus on test-based accountability by imposing sanctions on schools that did not meet set standards (Anderson, 2012). With the intensifying pressure for schools to demonstrate student performance, the term ‘high stakes’ testing came to be used when referring to the use of standardized tests to measure student success and the associated sanctions, or “stakes,” imposed on schools that did not meet benchmarks that indicated progress in student learning.

Given the importance of science education, it is critical to understand how education policy influences the instruction of science. Research suggests that test-based accountability policies influence the practice of science education in several ways. First, these policies compete with other, research-based reform efforts. For example, the AAAS, NRC, and NSF support the use of constructivist learning, inquiry-based instruction, student-centered teaching, and project based learning (Anderson, 2012). However, educators have indicated these inquiry practices are not utilized because of the excessive time consumed by efforts to improve standardized test scores (Goetz Shuler, Backman, & Olson, 2009; Kersaint, Borman, Lee, & Boydston, 2001). School administrators in particular have expressed difficulty balancing accountability policies with promoting constructivist approaches to science education (Kersaint et al., 2001). In addition, teachers have reported altering their instructional practices to accommodate what they perceive to be accountability mandates, expressing that state tests, not their professional opinion, became the main influence on instructional practices (Font-Rivera, 2003). Teachers also indicated feeling they must teach to these tests (Pinder, 2008), no longer teach the way they think is best for their students (Aronson, 2007; Galton, 2002) and reported inquiry-based lessons happening less frequently (Coble, 2006; Katzmann, 2007; Wideen, O'Shea, Pye, & Ivany, 1997).

In an attempt to synthesize empirical research findings on the relationship between science education and test-based accountability policies, Anderson (2012) analyzed 35 empirical studies dealing with test-based accountability and science education in order to examine the relationship between the two. Findings from Anderson's research indicate that test-based accountability policies correlated with changes in instructional practice, the amount of science taught, and teacher satisfaction. Specifically, administrators and teachers expressed feeling that accountability based reform limits the amount of time and effort spent on science, drives

instruction toward memorization of facts, and constrains teachers' use of professional judgment and creativity. Additionally, teachers reported experiencing increased stress, feeling their professional judgment was being constrained, and having fewer avenues for teaching creatively when having to standardize their behavior.

Measurement of Student Science Literacy

Wenning (2006) claimed that given the importance placed on achieving widespread science literacy, it would seem likely that some means of assessing students' progress toward that goal should exist. The author argued there is no such instrument, stating:

> The failure to have an instrument for assessing science literacy is probably due to several major reasons: (1) definitions of science literacy can incorporate a wide range of types, dimensions, and degrees; (2) a definition of science literacy will necessarily be complex if it is to be comprehensive and therefore meaningful; (3) a comprehensive assessment instrument would be of unacceptable length; (4) no single "high stakes" assessment instrument could provide all the information needed by teachers, school administrators, and agencies to make decisions to improve student learning; (5) there appears to be a confusion about educational purpose, teaching methods, and student outcomes, and (6) no one speaks officially on behalf of the world of scientists, philosophers, and educators who can advance by fiat a universal definition of science literacy. (Wenning, 2006 pp. 4-5)

Wenning (2006) also contends that none of the science education reform efforts to date have resulted in a significant attempt to assess the degree of science literacy of students and that standards-based educational reform has resulted in 'competency' tests which are being given on state, national, and international levels as demonstrated by such programs as state-mandated No

Child Left Behind (NCLB) testing, the National Assessment of Education Progress (NAEP), and the periodic Trends in International Mathematics and Science Survey (TIMSS). These tests, according to Wenning (2006), are achievement tests and not oriented toward assessing science literacy in a authentic way.

Methodological Approaches

Given the different conceptualizations of science literacy and the evolution of science reform efforts, it is not surprising that there are also variations in the manner in which science literacy is measured. Laugksch (2000) asserted that, although science literacy is mostly regarded as being of importance for general education, there are at least three ‘interest groups’ involved in science literacy, “namely (a) sociologists of science or science educators with a sociological approach to science literacy; (b) social scientists and public opinion researchers; and (c) science educators.” (p.87). Each interest group takes a different approach to measuring science literacy as evidenced from the methodologies utilized by the different groups.

The sociological approach to investigating science literacy attempts to identify and describe the range of possible interactions between people’s existing understandings of situations involving science and the understandings that originate from science itself (Wynne, 1991). In order to describe the science literacy of adults, this approach utilizes qualitative, contextual, small-scale, interpretive studies rather than large-scale samples and standardized questions. The main methods of obtaining data for this approach are case studies involving participant observation, longitudinal interviews, structured in-depth interviews, and questionnaires on specific local issues (Wynne, 1991). Wynne (2001) contends it is important to recognize that people assess whether or not they can use or trust expert knowledge partly by comparing it against elements of their own already-tested knowledge and experience. Wynne’s research

attempts to locate issues of the public understanding of science within specific practical social contexts. Examples of this type of research include a study of individuals who have inherited a gene that raises blood cholesterol levels, an investigation into individuals living in two communities close to hazardous industry, and an analysis of the role played by scientists in environmental organizations (Wynne, 2001).

The approach taken by public opinion researchers in measuring science literacy is significantly different from the sociological approach and is aimed at describing and comparing trends in the acquisition of specific scientific content knowledge, attitudes toward science, and support for science among a representative sample of the population. These researchers use large-scale samples, standardized questions, and survey techniques to obtain data. The work of Miller has been influential within this framework (Laugksch, 2000). Miller's 'three constitutive dimension' model of science literacy allowed the construct to be measured in this manner because his definition of science literacy is sufficiently specific and bounded. Advocates of the sociological approach to measuring science literacy have criticized the research methods of the public opinion model, naming it 'deficit model' (Laugksch, 2000). Ziman (1991) claimed the deficit model attempts to interpret science knowledge held by individuals in terms of public ignorance, or what they do not know, and does not provide an adequate framework for many of the results of research. Liu (2009) further argued that a deficit model ignores the fact that students and the general public possess a wide variety of informal knowledge and experiences about natural and life phenomena. And while informal knowledge and experiences may not be compatible with the commonly accepted scientific views, they are 'functional' in everyday contexts because they seem to explain various phenomena to their own satisfaction.

Science educators have taken yet a different approach to measuring K-12 science literacy. Attempts to measure the science literacy of students have focused more on content knowledge. Laugksch (2000) argued that science education researchers have tended to focus on a single dimension, such as content knowledge or student attitudes toward science, and referred to them as measures of science literacy. The author contends that a number of instruments have been developed to investigate a single dimension of science literacy, but few comprehensive measures exist to simultaneously assess several different dimensions of science literacy.

Calls to Measure Student Science literacy as a Way of Thinking

Herbert Spencer (1884) offered an argument for the benefits of science in his classic essay *What Knowledge Is of Most Worth*? In terms of various everyday activities (from self-preservation to leisure activities) he claimed that each could be accomplished more effectively by understanding how the natural world operates (Rudolph, 2014). In Spencer's view, it was the facts pertaining to nature that were most useful. Educational leaders were attracted to Spencer's approach and he heavily influenced the early curricular focus on science as a body of information and facts that should be learned (Rudolph, 2014).

However, in 1910 John Dewey proposed an alternative focus and called for a reconceptualization of what science education should aim to accomplish. Dewey claimed the tendency of science teachers is to present the subject of science as, "an accumulation of ready-made material with which students are to be made familiar" and "not enough as a method of thinking" (Dewey, 1910, p.122).

Dewey further stated:

> The responsibility of science cannot be fulfilled by educational methods that are chiefly concerned with the self-perpetuation of specialized science to the neglect of influencing the much larger number to adopt into the very makeup of their minds of those attitudes of

> open-mindedness, intellectual integrity, observation and interest in testing their opinions and beliefs that are characteristic of the scientific attitude… As long as acquisition of items of information, whether they be particular facts or broad generalizations, is the chief concern of instruction, the appropriation of method into the working constitution of personality will continue to come off a bad second (Dewey, 1934. p. 4).

Dewey recommended doing away with the overemphasis on "subject matter" and re-envisioning science education so that an understanding of the method of science was the desired outcome. Dewey believed that science education should teach what he referred to as "the scientific habit of mind" and provide students with a more complex understanding of science, giving them the ability to engage in empirical reasoning within their everyday lives (1910, p.126). Following Dewey's appeal, many science educators began to think about the formal definition and measurement of science literacy (Miller, 1983). Additionally, current learning theories recognized the importance of both formal and informal education as evidenced by an ever growing body of literature on student learning that shows students learn science in informal settings outside of school just as much as they do inside schools (Falk, 2001; Martin, 2004). For example, students engage with science every day through activities such as watching television and visiting museums or parks (Liu, 2009).

Liu states:

> Science literacy should be an evolving state instead of a status to acquire. People constantly learn science in and outside school, within and outside work, and both formally and informally. Learning science is indeed a life-long process, rather than the goal to achieve once and for all (p. 306).

Liu (2009) further suggests the unsatisfactory state of science literacy within the U.S. may be due to our outdated notion of science literacy and that broadening the notion of science literacy by including both extrinsic and intrinsic aspects and considering science literacy as a lifelong process is more in line with current views of how people learn. Others called for broader conceptualization of science literacy as well. For example, DeBoer (2000) contends "instead of defining science literacy in terms of specifically prescribed learning outcomes, science literacy should be conceptualized broadly enough for local school districts and individual classroom teachers to pursue the goals that are most suitable for their particular situations" (p. 582).

The National Research Council (NRC) has adopted the position:

> At its core, scientific inquiry is the same in all fields. Scientific research, whether in education, physics, anthropology, molecular biology, or economics, is a continual process of rigorous reasoning supported by a dynamic interplay among methods, theories, and findings. It builds understanding in the form of models or theories that can be tested. (Scientific Research in Education; National Research Council, 2002, p. 2).

Despite advocacy for measuring student science literacy beyond memorizing factual information, Rudolph (2014) claimed that schools still focus almost exclusively on content, rarely diverting from the goal of presenting as many facts as possible for the purpose of increasing standardized test scores. Jenkins (2003) also contends that items on existing measures of science literacy are mostly information-dependent, producing the need to focus on content knowledge in order for students to respond accurately. In contrast, Fives et al. (2014) argued that science literacy should emphasize those aspects of science that transcend specific fields/disciplines and focus on the processes of science. The National Science Board (NSB) has

echoed this sentiment in calling for the improvement of science assessments to demonstrate students' ability to think and apply knowledge by developing tests that measure both the subject knowledge and overall student understanding of science concepts. The NSB further recommended assessments that measure problem-solving skills and support learning that enhances the application of knowledge (National Science Board, 2006).

Under current test-based accountability policies, many believe that 'what gets measured gets taught.' However, given the fundamental importance of science, there is a need to assess science skills that involve critical and analytical thinking and not just simple recall. Layton, Davey, and Jenkins (1986) point out that most assessments of science literacy are based on a set of decontextualized items that are selected arbitrarily. This type of assessment results in an unreliable connection between a correct response to a test question and conclusions about an individual's ability to fully understand and engage intelligently with science. In other words, the scientific knowledge being tested does not connect well with the authentic situations in which students might use it. Feinstein (2011) stresses that, historically, making science relevant has meant presenting students with facts and principles. The author suggests a reconceptualization of instruction that moves away from 'making science relevant' as something the teacher does, to something that students learn to do, and increase proficiency through practice. This type of instruction would require students to start with their own questions, based within their own social context, which they then join with science in an attempt to connect scientific ideas with lived experience. Feinstein asserted that most students currently sit through a long series of presentations that include concepts and theories, with few students having the opportunity to test their own hypotheses, present their own results, and critique each other's findings.

Expanding on this perspective, Feinstein, Allen, & Jenkins (2013) assert that schools should help students identify and interpret scientific information needed to face practical problems as well as evaluate the credibility of the information based on evidence. The authors state that while some people are interested in science for its own sake, most individuals utilize science in situation-specific contexts, using it to help them solve problems. To illustrate this point, Feinstein (2014) points to the example of a mother seeking treatment for her autistic son, claiming that while she may seek out research literature, she is not attempting to understand that literature from the perspective of a scientist. Rather she is attempting to incorporate what she learns from the literature with her first-hand knowledge of local services and an understanding of the needs of her child. Feinstein further states that it is important to be realistic about the types of scientific understanding people need in order to make important life decisions. He suggests a reconsidering of the goals of science education, insisting that pedagogy should rely on students to identify scientific information that is useful to them rather than focusing of the general principles of science.

Student Attitudes Toward Science

Recent reports on pedagogical approaches to science frequently reflect the opinion that the content being taught in classrooms is not aligned with the interests of most students or the needs of society (NSB, 2010). Critics of traditional teaching practices claim that students do not find their science classes interesting and motivating because the curricula is overloaded with content that emphasizes isolated facts that are detached from practical applications. As a result, students do not make connections between the facts presented in class and the relevance to their lives. This perceived lack of significance leads to low levels of motivation and generally diminished interest in science (Hofstein, Eilks, & Bybee, 2011).

Recognizing a lack of interest in science is important because student attitudes have been linked to motivation, learning, and behavior (Osborne, Simon, & Colling, 2003). In fact, a substantial body of literature accumulated which associates academic achievement with attitudinal and affective variables such as self-concept, confidence in learning, science interest, motivation, and self-efficacy (Singh, Granville, & Dika, 2002). These factors also predict science avoidance, which in turn affects student's long-term achievement and career aspirations in the science field (Reynolds & Walberg, 1992).

Early inspiration for attitude research in science education stemmed from John Dewey's (1916) philosophy, which underscored the need for teaching scientific attitudes as an important aspect of education. He also argued that science instruction should promote intellectual integrity and open-mindedness rather than simply communicating a fixed body information (Dewey, 1934). Research conducted since Dewey's writings on educational reform has, in fact, shown that science learning cannot be explained by examining cognitive factors alone (Koballa, 2013). Rather, students' beliefs towards science are as predictive of achievement as previous grades or standardized test scores (Chen & Pajares, 2010). Furthermore, findings related to student attitudes toward science extend beyond a specific field or course. Research suggests that a deep, personal interest in any field of science provides motivation to interact with other scientific fields in the future. Also, students who pursue their own science related interests have an increased confidence in their ability to learn science in the future and are less likely to lose interest over time (Feinstein, Allen, & Jenkins, 2013).

Student beliefs play a particularly important role during early adolescent years and during periods of transition, such as from elementary school to middle school (Chen & Pajares, 2010). According to the National Science Educational Standards, teachers should "select science

content and adapt and design curricula to meet the interests, knowledge, understanding, abilities, and experiences of students" (NRC, 1996, p. 30). The NRC further claimed that "[s]cience literacy begins with attitudes and values established in the earliest years…" (p.18) and "… attitudes and values established toward science in the early years will shape a person's development of science literacy as an adult" (p.22).

Science literacy frameworks developed more recently have included student attitudes toward towards science. In developing an instrument to measure the science literacy of middle school students, Fives et.al (2014) reviewed research literature on student's motivation and beliefs in science. The authors selected three constructs relevant to successful student engagement in science for inclusion in their instrument: self-efficacy, subjective task value, and personal epistemology.

Self-efficacy

Self-efficacy as it relates to science refers to a student's beliefs about his or her personal competence to succeed in science (Zusho & Pintrich, 2003; Joo, Bong, & Choi, 2000). Research has found that students who have strong efficacy beliefs toward science are more likely to engage in such tasks, work hard to complete them successfully, and persist when faced with obstacles. Conversely, students who do not believe that they can succeed in science-related activities will avoid them, apply minimal effort when engaging in science tasks, and be more likely to give up when faced with challenges (Britner & Pajares, 2006).

In order to expand previous research regarding the influence of student self-efficacy within areas such as math and language arts to the area of science at the middle school level, Britner & Pajares (2001) conducted a study of 272 seventh grade students to examine whether the science motivation beliefs of middle school students vary as a function of their gender or

race/ethnicity. The study also sought to determine whether science self-efficacy beliefs predict science achievement when motivation variables shown to predict achievement in other academic areas are controlled. Findings indicated that girls reported stronger science self-efficacy and earned higher grades in science compared to male participants. White students had stronger self-efficacy and achievement than African American students and self-efficacy was the only motivation variable to predict science achievement. The authors suggest findings of this study strengthen Bandura's (1986) contention that self-efficacy beliefs play an influential role in human agency and supports prior findings of a significant connection between self-efficacy beliefs and related academic outcomes. The authors further recommend that school practitioners should examine students' beliefs about their academic capabilities as important predictors of other affective variables and academic performances and efforts should be made to identify and nurture these beliefs (Britner & Pajares, 2001).

Task Value

One of the most prominent theoretical perspectives on the development of academic beliefs is the expectancy-value model. First developed by Atkinson in 1957, the model theorizes achievement-related behaviors, such as selecting a task and persisting in it, can be explained by students' expectations for success and value (Kahraman & Sungur-Vural, 2014). Building on the work of Atkinson, a more contemporary theory of expectancy-value has been developed which suggests that achievement related choices are influenced most directly by subjective task values (intrinsic, utility, and attainment values, and costs) and expectancies of success (Eccles and Wigfield, 2002; Pintrich and Schunk, 2002). Within the more modern theory, attainment value refers to the importance students' place on doing well, student's interest in the task is referred to as intrinsic value, utility value relates to the task's level of importance for the students' current

and future goals, and cost can be defined as negative consequences associated with engaging in the task (Eccles and Wigfield, 2002; Pintrich and Schunk, 2002; Eccles, 2009).

Students' perceived self-efficacy for a particular task has been shown to relate to the value they place on that task (Zepeda, Richey, Ronevich, & Nokes Malach, 2015). In a longitudinal study of 761 students in grades 1 through 12, Jacobs, et al. (2002) examined the effect of changes in perceived self-competence on subjective task-value within the same domain. Results indicated that self-perceptions of competence and subjective task values declined as children got older and changes in students' competence beliefs were strongly associated with changes in value for an activity. Specifically, declines in perceptions of ability accounted for over 40% of the decrease in task value. The authors assert that competence beliefs and task value should be studied together.

Furthermore, the research of Zepeda, Richey, Ronevich, and Nokes-Malach (2015) indicates that instructional intervention that improves metacognitive skills can improve the motivational aspect of learning. In order to test whether an instructional intervention which attempts to improve metacognitive skills results in changes to motivation, metacognition, and learning, the researchers conducted an experiment in which 46 eighth-grade students were assigned to a treatment group, which received problem-solving practice along with metacognitive instruction and training, or a control group, which received extensive problem-solving practice. Results indicated that students in the metacognitive intervention reported greater task value for the science material covered in class than the control group. The authors suggest that the intervention materials may have helped students to recognize the value of learning the course material.

Student Epistemological Beliefs

Epistemology is a branch of philosophy that relates to the origin, nature, validation, and limitations of knowledge (Hofer, 2002). Epistemic beliefs in science, or the ways in which a person thinks about how science knowledge is gained and the sources of that knowledge, can guide a person's basic focus toward scientific information and how they evaluate scientific information (Sinatra, Kienhues, & Hoffer, 2014). Studies have found that more sophisticated epistemological beliefs are associated with higher levels of achievement and learning (Hofer & pintrich, 1997, Schommer, 1993). Schommer (1993) investigated the development of over 1,000 high school students' epistemological beliefs and the influence these beliefs have on academic performance. The study was designed to assess four beliefs; simple knowledge, the belief that knowledge is best described as isolated facts; certain knowledge, knowledge is absolute; innate ability; the ability to learn is innate; and quick learning, gaining knowledge happens quickly or not-at-all. Results indicated that epistemological beliefs predicted students' grade point average. Specifically, the less students subscribed to beliefs in simple knowledge, quick learning, certain knowledge, and fixed ability, the higher their grade point average.

Similarly, Ricco, Pierce, and Medinilla (2010) studied the influence of epistemic beliefs and motional variables to the prediction of student grades and self-regulation in 459 predominantly Hispanic, lower-income middle school students. Results indicated that epistemic beliefs predicted science grades more than motivational factors. Students who perceived the knowledge development as changing with new discoveries tended to have higher grades, whereas, students with a viewed learning as an automatic process that is based on ability had lower grades.

Evidence suggests that certain types of science instruction, hands-on classroom

experiences that involve students in the design of experiments and the collection and analysis of data, may promote epistemological thinking (Solomon et al., 1996). During a year-long classroom study, Solomon, Duveen, and Scott (1994) interviewed and administered surveys to approximately 300 students, ages 11 to 14 years old, about the purpose of experiments, and the nature of scientific theories. They also asked students to provide an example of an experiment and explain how it helped them to understand a theory. The authors found it was rare, especially for younger students, to think of scientific experiments as a purposeful activity conducted with the goal of generating and testing explanations. Students were also not able to differentiate between descriptions and explanations. In a subsequent large-scale study, in which Solomon and colleagues (Solomon, Scott, & Duveen, 1996) used the same questionnaire with a much larger age-range of students (1,000 students, ages 12- to 18-years-old), the researchers found that older students' displayed significant progression toward a more sophisticated understanding of science. However, only 20 percent of the students could correctly describe the theory related to a particular experiment. Elder (2002) found that students' epistemological beliefs in science reflected both mature and naïve understandings. In particular, students displayed sophisticated understandings of scientific knowledge, reflecting the idea that knowledge comes from reasoning, thinking, and experimentation. However, in their interviews, students also indicated a less sophisticated belief that the purpose of science was to do projects and activities, rather than explain phenomena. Students also saw their role in scientific activities as passive, with most students reporting sources such as books, teachers, or family members as the source of scientific ideas. However, when asked about the knowledge acquisition of scientists, students perceived scientists as having more active roles, indicating scientists' ideas originate from curiosity, exploration, or interactions with the environment.

Science educators and researchers have remarked that most science education is not informed by epistemology, suggesting that typical lessons require students to engage in recipe-like labs or activities without an understanding of the purpose of those activities. Smith et al. (2000) interviewed two groups of 6th grade students in order to investigate whether elementary school students could make significant progress in developing more sophisticated, constructivist epistemology, when provided with a science curriculum designed to support students' epistemological thinking. The first group, comprised of 18 students, was taught from a constructivist perspective in which students actively participated in the development, testing, reflection, and revision of their own ideas about how things work through collaborative inquiry with other classmates. The control group, comprised of 27 students, was taught using traditional instructional methods. Results indicated that the elementary students were able to construct more sophisticated epistemological views than had been believed. Students in the more traditional classroom environment developed an epistemological view of science as acquiring factual knowledge or involving simple activities and procedures. This was characterized by beliefs that gaining knowledge occurs in a piecemeal fashion by making observations and doing experiments without differentiating between ideas, experimental methods, and results. Conversely, students in the constructivist classroom developed an epistemological view that focused on the role of ideas in the process of gaining knowledge and on the mental, social, and experimental work involved in understanding, developing, testing, and revising ideas.

Measuring Student Attitudes Toward Science

The call for re-conceptualizing science instruction necessitates changes to science assessments. Advocates of teaching methods that promote affective elements of science learning have urged policy makers, state departments of education, and local schools to specifically

address affective elements of learning in their science curricula and accompanying assessment programs. The rational being attitudes toward science play an important role in science literacy and personally fulfilling learning experiences are more likely to result in positive attitudes and increased motivation in science learning and achievement (Osborne, Simon, & Colling, 2003).

An important goal of science education is to help students develop interest in and support for scientific inquiry as well as to obtain and apply scientific knowledge for the benefit of themselves and society (Bybee, McCrae, & Laurie, 2009). However, few measures of student attitudes toward science instruments have sufficient psychometric data to serve as a reliable source of information (Osborne, Simon, & Collins, 2003). Blaylock et al. (2008) conducted a comprehensive review of 66 science attitude instruments which were grouped into the following categories: attitudes toward science, scientific attitudes, nature of science, scientific career interests, and other. The authors found evidence that the field lacks instruments which are both valid and reliable. Instruments were evaluated against five criteria: the extent to which they were theoretically grounded; tests reliability; measures used to establish validity; how the dimensionality of the instrument had been used in reporting scores; and the extent to which the instrument had been tested and developed prior to its use. Of the 66 attitude instruments evaluated, 14 offered no reliability or validity evidence, 12 gave reliability evidence but no validity evidence, and two had validity evidence but no reliability. Therefore, 28 of the 66 instruments (42%) were missing some fundamental psychometric component. The authors expressed concern with the general quality of the studies because of lack of planning or thoughtfulness in design, and claim that some drew conclusions far beyond the limits of the data. Additionally, few studies considered the effects of missing data, and 37 of the 66 instruments (56%) were used in a single study with no published follow-up studies.

While there is value in assessing students' understanding of field specific science concepts, one could argue that measuring students' ability to understand science as an approach is even more important, given our modern societal emphasis on science and technology. However, assessing scientific literacy from this perspective cannot be achieved without valid and reliable measures that allow educators and researchers to make appropriate inferences about students' ability to think scientifically. Such measures could serve as a valuable formative assessment tool within the classroom, as well as a method for researchers to add to the body of knowledge surrounding instruction.

Summary

This chapter provided an overview of key literature regarding the conceptualizations and definitions of science literacy as well as a summary of science educational reform in the United States. The literature provides evidence of the important implications for how evolving perspectives regarding the construct have shaped policy, which in turn impact classroom practice. Much of the literature calls for the measurement of science literacy as a way of knowing rather than placing an emphasis on requiring students to memorize decontextualized, discipline specific content, thus broadening the notion of science literacy to an ongoing, lifelong process.

A variety of studies have also linked student attitudes with learning, motivation, and behavior. Within the field of attitude research, student attitudes and beliefs have been shown to play a particularly important role during early adolescent years. Additional studies have indicated that classroom practices designed to promote affective and epistemological beliefs can result in students' engaging more actively with science and obtaining more sophisticated knowledge.

Finally the literature regarding science literacy emphasizes a current need to explore assessing science literacy in a more comprehensive manner in order to align instruction to student interests and the ways students engage with science on a daily basis. However, the field lacks instruments possessing the psychometric qualities required to serve as reliable sources of information. This literature points to the current need to expand the definition of science literacy to include student attitudes and beliefs as well as develop valid and reliable measures to assess these affective elements.

Chapter 3

Research Methodology

According to the *Standards for Educational and Psychological Testing* (APA, 2013) validity is the most fundamental consideration in developing and evaluating instruments. Validity refers to the degree to which accumulated evidence supports the meaningful interpretation of test scores for proposed uses of a test (Creswell, 2013). This concept of validity applies not only to tests, but also other types of measures such as attitudinal surveys or scales. The *Standards* put forth five sources of evidence for evaluating the validity of interpreting test scores for a particular use; these include evidence based on test content, response process, internal structure, relations to other variables, and consequences of testing (APA, 2013).

The developers of the Scientific Literacy Assessment (SLA) examined validity evidence based on test content, response process, and internal structure (Fives, Huebner, Birnbaum, & Nicolich, 2014). Specifically, test content validity was addressed using the following strategies: (1) item development specifically aligned with each of the components identified in the theoretical framework; (2) input from an interdisciplinary research team composed of two epidemiologists, a statistician, and an educational psychologist who had also been a middle school science teacher for six years; (3) evaluation of initial measure by experts in science education; and, (4) utilizing a previously validated process for creating items. Evidence based on response process was provided by conducting think-aloud interviews with six middle school students during two pilot studies conducted during development of the instrument. Evidence based on internal structure was examined by conducting principal components factor analysis on the 25 item SLA-MB as well as the 26 item SLA-D. While the SLA-D was found to be

unidimensional, results indicated the SLA-MB assesses three distinct set of beliefs about science literacy. Reliability analyses also provided evidence of internal consistency. The Kuder-Richardson equation 20 (1937) indicated good reliability for both versions of the SLA-D, producing coefficients of0.83 and 0.82. Cronbach's alpha was computed for each of the three SLA-MB subscales. Each demonstrated sound reliability (value of science: α =0.80, self-efficacy for science: α = 0.72, and source and certainty of scientific knowledge: α = 0.88).

While Fives et al. provided evidence of internal structure and reliability, the authors considered the SLA to be a 'work in progress' and invited others to participate in the use and evaluation of the instrument in order to provide a wider range of validity evidence through a process similar to 'crowd sourcing' (p. 576, 2014). During the initial development process, a shortened version of the SLA-D containing 19 items was created, however validity evidence for the shorter version was not explored. This research sought to expand the developers' initial validation research by examining validity evidence for the 19-item version and to extend the initial work of Fives et al. by examining additional types of validity evidence. Specifically, this research examined evidence based on internal structure and relationships to other variables by utilizing existing survey data that were collected during the research efforts of Project CRESST: Enhancing Clinical Research Education for Science Students and Teachers,[1] a curriculum and professional development program for middle and high school science and health/physical education (HPE) teachers. Project CRESST was designed to increase awareness of the clinical research process and support the instruction of research concepts and skills using inquiry-based instructional approaches in middle and high school science classrooms. Funded by the National Institutes of Health, CRESST provides teachers with an intensive professional development

[1] Project CRESST was supported by the National Institutes of Health, Office of the Director, Science Education Partnership Award through Grant R25 OD010984-05 to the Virginia Commonwealth University.

experience and classroom lessons and activities to support inquiry-based teaching of science content and skills. In order to participate, teachers went through a competitive application process and were selected from a pool of potential candidates. Each summer between 2011 and 2015, CRESST hosted an intensive one-week summer workshop, where teachers were provided with interactive experiences and curricular material specifically designed to increase participants' understanding of clinical research and facilitate inquiry based science instruction. Teachers also interacted with university scientists and individuals conducting ongoing clinical research studies. Teachers continued their professional development throughout the fall semester by utilizing the project curricula and tools in guiding their students in research projects and sharing their experiences with other academy colleagues. Teachers earned three graduate credits or were awarded a stipend of $500 for participating in the academy and follow-up activities. In total 93 teachers from 64 schools throughout Virginia attended the CRESST Professional Development Academies. Following the success of the summer academies, the Project converted to an online format in 2016, providing teachers greater flexibility to participate in the program.

Program evaluation and research efforts for Project CRESST were designed to assess participants' self-efficacy and confidence related to instructional practices when teaching research concepts. Data were collected to examine the impact of teacher participation on instruction and student learning. In 2014, research efforts were expanded to include a focus on science literacy to obtain baseline measures of teacher and student's reported levels of science literacy. As part of these efforts, teachers from Academies 1 through 4 (2011-2014) were recruited for participation. This data collection included the administration of surveys to measure students' science literacy, content knowledge, and attitudes toward science. The current

research utilized this existing data to conduct a validation study of SLA.

Research Design

This investigation involved a secondary analysis of survey data. In order to address the research questions, a quantitative nonexperimental correlational design was used. Correlational research involves the systematic investigation of the associations among variables and provides empirical evidence to describe and measure the relationship, or lack of relationship, between two or more variables or sets of scores (Creswell, 2013). While this evidence does not establish causal relationships, it does contribute to a deeper understanding of the variables and serves as a powerful tool for examining sources of validity evidence including: convergent, discriminant, criterion-related, concurrent, and predictive validity (Cunnington, Weathington, & Pittenger, 2013). This research utilized two of the most common correlational design types, descriptive and predictive designs. As the names suggest, descriptive correlational methods describe the variables and the relationships that occur naturally between and among them while predictive correlational studies predict the variance of one or more variables based on the variance of another variable(s). Within predictive correlational designs, as with experimental designs, the variables being studied are labeled as either independent (predictor) or dependent (outcome) variables. However, unlike experimental designs, the predictor variables within correlational designs are not manipulated, but occur naturally (Sousa, Driessnack, & Menendes, 2007).

Table 1 provides an overview of the study and includes the specified research questions, aligned sources of validity evidence, and data analytic procedures.

Table 1 *Research Questions, Sources of Validity Evidence, & Analytic Procedures*

Research Question	**Type of Validity Evidence**	**Analysis Procedure(s)**
1. To what degree is the SLA a valid measure of middle school students' general scientific knowledge, motivation, and beliefs related to science epistemology?	Evidence based on internal structure	- SLA-MB – exploratory factory analysis - SLA-D1, 19 item version – exploratory factor analysis - Cronbach's alpha
2. To what degree is the SLA a valid measure of high school students' general scientific knowledge, motivation, and beliefs related to science epistemology?	Evidence based on internal structure	- SLA-MB – exploratory factory analysis - SLA-D1, 19 item version – exploratory factor analysis - Cronbach's alpha
3. To what extent do scores on the SLA-MB correlate with scores on The Modified Attitudes Toward Science Inventory (mATSI) and the STEM Career Interest Survey (STEM-CIS)?	Concurrent criterion-related validity evidence	- Pearson Product-moment Correlations - Cronbach's alpha
4. To what degree do scores on the SLA-D predict scores on a student general science knowledge test?	Predictive criterion-related evidence	- Bivariate Linear Regression - Cronbach's alpha

Recruitment

Recruitment of participants for the Project CRESST study component utilized within this research began in May of 2014 following the approval of Virginia Commonwealth University's Institutional Review Board (IRB). Recruitment emails were sent to 38 middle and high school science teachers from cohorts one through four of Project CRESST's annual professional development academies. Because this assessment was aimed at evaluating the instructional and curricular impact of the science component of the program, health and physical education teachers were excluded from recruitment efforts. The recruitment correspondence summarized activities associated with participation, which included administering survey packets to students within their classrooms during two data collection sessions, with the first set of instruments being administered during fall 2014 and the second during spring 2015. After sending the initial recruitment correspondence, two follow up emails were sent over a four-week period. As an incentive to participate, teachers were paid $50 for each administration of the surveys. Of the 38 teachers invited to participate, five responded resulting in a 13% response rate. Participating teachers were provided with a standard active consent form prior to data collection.

Participants

Participants were five teachers who attended one of the Project CRESST Summer Academies. All members of this convenience sample were White females and represented a range of educational settings. Two of the teacher participants were based in rural middle schools and taught Life Science classes. The third middle school teacher taught Physical Science within a suburban school. One high school teacher participant taught Biology within a rural high school. The second high school teacher was based in a suburban school and taught a Health Science Chemistry class as part of a Health Science Specialty program, designed to prepare students for careers in the healthcare industry. Survey packets were completed by 400 students

enrolled within 20 general education science classes taught by the five teachers. Overall, 60% of the students were female, 66% were white, and 65% of students were in middle school. Detailed descriptive statistics for the student sample can be found in Table 2. Compared to the state population, female students were over represented in the study, comprising 60% of respondents as compared to the 49% of female students in the state. White students were also over represented in the sample, 66% as compared to comprising 51% of the state student population (Virginia Department of Education, 2017).

Table 2 *Characteristics of Students Completing Survey Packets (N=400)*

	n	%
Gender		
Male	161	40
Female	239	60
Race/Ethnicity		
White	262	66
Black/African American	79	20
Hispanic	31	8
American Indian	13	3
Other	23	6
Grade		
7	155	39
8	103	26
10	142	36
Age		
12	105	26
13	106	27
14	47	12
15	82	21
16	60	15
Subject		
Biology	70	18
Health Science	72	18
Life Science	155	39
Physical Science	103	26

Original Data Collection

There were two stages of data collection, the first in fall 2014 and the second during spring 2015. Two weeks prior to the first stage of data collection, each teacher distributed a Student/Parent Information Form explaining the purpose of the study, a description of students' involvement in the study, efforts to ensure student confidentiality, and the voluntary nature of participation. Parents were also informed about the option to request their child be excluded from participation. Prior to completing the assessments, the researcher distributed and reviewed the student assent form. Once the instrument packets were distributed, students created their own unique identification code that allows for the linking of the spring and fall science instruments. The codes were known only to the students and were generated by responses to a series of questions. The only identifying information collected for students was a signed consent form which was collected separately from the instruments. The first stage of data collection, occurring during fall 2014, contained a questionnaire to obtain student demographic information as well as measures of students' science literacy concepts, general scientific knowledge, motivation and beliefs related to science, and attitudes towards science and interest in science careers. The order of the measures administered during the first stage of data collection was varied by class in order to mitigate order effects that might result from completing the assessments. During the second stage of data collection, during spring 2015, students completed an instrument designed to assess their science content knowledge. In total, data were collected within 20 science classes, resulting in the administration of survey packets to 400 students.

Measurement

Assessment of student science literacy. The Scientific Literacy Assessment (SLA) developed by Fives, Huebner, Birnbaum and Nicolich (2014) was utilized as a measure of students' science literacy as well as their motivation and beliefs related to science epistemology.

The SLA is comprised of two parts. The SLA-D (See Appendix A) assesses five components of demonstrated scientific literacy: role of science, scientific thinking and doing, science and society, science media literacy, and mathematics in science. There are two versions of the demonstrated component, SLA-D1 and SLA-D2, each includes 26 selected response items (11 shared items and 15 unique items). Each version of the SLA-D exhibits good reliability estimates, producing Kruder-Richardson equation 20 scores 0.83 and 0.82 respectively. The second component of the SLA measures student scientific literacy motivation and beliefs (SLA-MB) and is comprised of three subscales: Value of Science (6 items), Scientific Literacy Self-efficacy (8 items), and Personal Epistemology (11 items). Respondents select from a five point response scale ranging from "Strongly Disagree" to "Strongly Agree." The SLA-MB (See Appendix B) also exhibits sound reliability estimates: Value of Science, α =.80; Self-efficacy for Science Literacy, α = 0.72; Source and Certainty of Scientific Knowledge, α = 0.88. Data for the initial validation of the SLA were obtained in May and June of 2012 within five Northern New Jersey schools. The 321 participants were in seventh or eighth grade. Female students were slightly overrepresented (56%) and participants reported ethnicities including Hispanic (34%), White (24%), Asian (21%), African American (14%), and other or multiple (7%).

Assessment of student attitudes toward science. The Modified Attitudes Toward Science Inventory (mATSI) developed by Weinburgh and Steele (2000), was utilized to assess student attitudes toward science. The mATSI (See Appendix C) is a 25-item, Likert-type instrument consisting of five scales: perceptions of the science teacher, anxiety toward science, value of science to society, self confidence in science, and desire to do science. Students respond to forced-choice responses ranging from "strongly agree" to "strongly disagree." Each statement is scored on a scale ranging from 6 for strongly agree to 1 for strongly disagree, and

summed across each scale. Higher scores represent more positive attitudes on each scale, except for the anxiety scale. Lower anxiety scores represent more positive attitudes. The mATSI exhibits good reliability estimates, producing an overall Cronbach's alpha of 0.70, and reliability coefficients for the five scales range from 0.72 to 0.78.

Assessment of student interest in science careers. The STEM Career Interest Survey (STEM-CIS), developed by Kier, Blanchard, Osborne, and Albert (2014) was utilized to assess student levels of interest in science careers. The STEM-CIS (See Appendix D) is an instrument that contains four discipline specific subscales, allowing for student career interest in science, technology, engineering, and mathematics to be measured separately or collectively. The Science subscale is comprised of 11 items (e.g. "I plan to use science in my future career" and "If I do well in science classes, it will help me in my future career"). Respondents select from a five point Likert- type response scale ranging from "Strongly Disagree" to "Strongly Agree." Each statement is scored on a scale ranging from 5 for strongly agree to 1 for strongly disagree, and summed across the statements for each scale. Higher scores represent more positive attitudes toward student interest in science careers. The Science subscale exhibits sound reliability producing Cronbach's alpha of 0.77.

Assessment of student science general knowledge. Project CRESST researchers sought to assess the general science knowledge of students as part of the original data collection. For that purpose, two sets of questions were assembled, one for middle school students and one for high school students, utilizing released items from the Virginia Standards of Learning (SOL) as well as items from the National Assessment of Educational Progress (NAEP). Items were selected by grade and subject then mapped onto the Virginia Department of Education's Standards of Learning (see Tables 3 and 4). The resulting item pool was used to create a 25-item

multiple-response instrument for middle school students (See Appendix E), and a 17-item multiple-response instrument for high school students (See Appendix F). Selected items were then reviewed by a university staff member, licensed to teach general science and biology in the state of Virginia. Once the final items were selected, the instruments were sent to participating teachers who reviewed the items to ensure they were appropriate for the students within their classes. All participating teachers affirmed the items were appropriate for their students.

Table 3 *Mapping of Items for High School General Science Knowledge Survey*

High School General Science Knowledge Survey			
Item Number	**Correct Answer**	**SOL**	**Question Origin**
1	A	LS.3a	NAEP
2	C	ES.6a	NAEP
3	D	ES.9c	NAEP
4	D	6.6a	NAEP
5	B	PS.8a	NAEP
6	A	PS.9a	NAEP
7	C	PS.4c	NAEP
8	C	BIO.5d	VDOE
9	D	LS.1h	NAEP
10	C	PS.10c	NAEP
11	B	6.1h, LS.1h	NAEP
12	B	6.5a	NAEP
13	A	BIO.1c, LS.1g	VDOE
14	D	PS.1j, PS.2f	NAEP
15	A	PS.5b	NAEP
16	C	LS.1i	NAEP
17	A	6.1f, LS.1e, BIO.1f	VDOE

Table 4 *Mapping of Items for Middle School General Science Knowledge Survey*

Middle School General Science Knowledge Survey			
Item Number	**Correct Answer**	**SOL**	**Question Origin**
1	A	LS.2a,b	NAEP
2	C	PS.10b	NAEP
3	C	ES.7a,b	NAEP
4	A	LS.2a	NAEP
5	A	6.2e, PS.6a,b	NAEP
6	A	6.8a, ES.3b	NAEP
7	B	PS.8a	NAEP
8	A	LS.13a	NAEP
9	B	PS.4a,b	NAEP
10	B	6.6d	VDOE
11	D	LS.12d	VDOE
12	D	PS.9b	NAEP
13	A	ES.9b	NAEP
14	A	6.2e	VEOE
15	C	LS.4c	VDOE
16	C	PS.2d	NAEP
17	D	LS.6a	NAEP
18	B	PS.9b	NAEP
19	B	PS.10b	NAEP
20	A	6.5b	NAEP
21	C	LS.8a	NAEP
22	C	LS.12e	NAEP
23	C	6.7a, LS.11e. ES.8e	NAEP
24	C	PS.10a,b	NAEP
25	B	6.8e	VDOE

Prescreening

Data were obtained from the principal investigator of Project CRESST in the form of two SPSS files, one for each data collection period. To ensure the accuracy of data entry, 81 (20%) of the survey packets completed by students were selected randomly for visual comparison to the electronic files. Within the 81 survey packets, only 4 data entry errors were found, suggesting the data had been entered with > 99% accuracy. SPSS was then utilized to examine descriptive statistics for all items. Frequencies, means, and ranges were inspected to ensure data were of the correct format and within the acceptable range for each set of response options. Once the accuracy of the electronic data files was confirmed, a series of statistical tests were conducted to ensure assumptions necessary for each analytical procedure were met.

Completeness. Prescreening for the presence and pattern of missing data is crucial for accurate data analysis (Tabachnick & Fidell, 2001). Commonly, researchers identify and categorize the degree of missing data as missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR) (Dattalo, 2013). Enders (2010) describes MAR as "ignorable missingness" (p.13) while data that are MNAR may result in bias and signal the potential for systematically missing responses. In order to identify patterns of missing data, missing responses were coded to a value of 1 and all other responses were recoded as 0. A series of bivariate correlation coefficients were then created to examine Pearson's *r* (*r*) to look for patterns. The resulting *r* values are classified as; less than 0.3 is considered as MCAR, 0.3 to 0.5 are considered MAR, and values over 0.5 are identified as NMAR (Dattalo, 2013).

In the current study, missing data was addressed on multiple levels by: (a) deleting cases with two or more missing values, (b) deleting cases identified as NMAR, and (c) the use of mean imputation for factor analysis. Where imputations were generated, bivariate correlations were

conducted between the means of the imputed (new variable) and non-imputed (original variable) to identify significant differences between the two. No significant differences were found, suggesting that the distributions of the variables had not been distorted by the imputation (see Table 5 for description of missing data).

Outliers. Cook's distance (D) was computed in order to provide a measure of the influence an outlying case has on the estimated regression coefficient and was obtained using the liner regression procedure within SPSS. Cutoff points for Cooks D were established for each instrument by taking the average of three commonly used methods. The first method employs the formula n/k, the second uses 4/(n-k-1), and the third is simply setting a cutoff value of 1 (Dattalo, 2013). For each of these methods, n equals the number of cases in the entire dataset and k equals the number of variables. For each instrument, Cook's D was computed on all items and compared to the calculated cutoff value.

Delimitations

The goal of this study was to expand upon the validity evidence of the SLA, developed by Fives, Huebner, Birnbaum and Nicolich (2014). Findings are delimited to the sample of middle and high school students who voluntarily completed the survey packets within the classrooms of participating teachers. However, the findings may be useful in supporting the meaningful interpretation of scores for this instrument.

Summary

This investigation utilized a quantitative nonexperimental correlational design utilizing survey research in order to expand upon the validity evidence of the Scientific Literacy Assessment (Fives et al., 2014). The survey data utilized were collected during the research efforts of Project CRESST: Enhancing Clinical Research Education for Science Students and Teachers, a curriculum and professional development program for middle and high school

science and health/physical education teachers. Statistical analysis included factor analysis to investigate the extent to which the SLA is an effective measure of science literacy for a high school population, the degree to which the 19-item version of the SLA-D is a valid measure of middle and high school students' general scientific knowledge, and the degree to which the SLA-MB is a valid measure of middle and high school students' motivation and beliefs related to science epistemology. Pearson's correlation coefficient was utilized to examine the extent to which scores on the SLA-MB correlate with scores on the Modified Attitudes Toward Science Inventory and the STEM Career Interest Survey to demonstrate concurrent criterion-related validity evidence. Bivariate regression analysis was utilized to investigate the degree to which scores on the SLA-D predicted scores on a student general science knowledge test, demonstrating validity evidence based on relations to other variables.

Table 5 *Description of Missing Responses and Cases Deleted Due to Missing Data*

Instrument/Number of Items	Number of Students who Completed the Instrument	Number/Percentage of Missing Items	Number/Percentage of Cases Deleted Due to Missing Data
SLA-MB (25 Items)	400	36 (0.3%)	2 (0.5%)
SLA-D - Middle School (19 Items)	258	62 (0.8%)	9 (2%)
SLA-D - High School (19 Items)	142	2 (0.001%)	0
STEM- CIS (11 Items)	378	13 (0.2%)	6 (1.5%)
mATSI (25 Items)	400	124 (1.2%)	5 (1.2%)
General Knowledge Instrument – Middle School (25 Items)	265	14 (0.2%)	9 (3%)
General Knowledge Instrument – High School (17 Items)	153	5 (0.2%)	0

Chapter 4

Findings

This chapter summarizes statistical findings from secondary analysis of survey instruments completed by a sample of middle and high school students. First are descriptions of factor analytic procedures conducted on the Motivation and Belief component of the SLA for both the middle and high school samples. Following are descriptions of factor analytic procedures conducted on the demonstrated scientific literacy component of the SLA for both the middle and high school samples. The next section details findings from Pearson Product-moment Correlations generated for the SLA-MB, MATSI, and STEM-CIS. The final analyses described include results from bivariate linear regressions conducted to explore the relationship between scores on the SLA-D and an instrument of student's general science knowledge for both the middle and high school sample. All analyses were conducted using the Statistical Package for Social Sciences (SPSS) version 24.

Characteristics of the Study Sample

Data for this research study were obtained from 259 middle school and 141 high school students, creating a total sample of 398 participants. Relationships among variables of interest (see Table 6) were then examined to address each research question. Within this sample, female students were over represented, comprising 60% of respondents as compared to making up 49% of the statewide student population. White students were also over represented in the sample, 66% as compared to comprising 51% of the state student population.

Independent samples *t* tests revealed significant differences between responses from the middle and high school samples on the STEM-CIS, SLA-MB, and mATSI (descriptive statistics for these instruments can be found in Tables 7 – 9). High school students expressed greater

interest in science careers as indicated by STEM-CIS scores (M = 41.54, SD = 7.328) than middle school students (M= 27.46, SD = 8.116), $t(370) = 4.82$, $p < 0.001$. High school students produced significantly higher scores on the SLA Value of Science Scale (M=22.10, SD=4.691) than middle school students (M=20.93 SD=5.008), $t(370) = 2.22$, $p = 0.027$ (see table 7 for item descriptions). Findings from the mATSI also indicated high school students in the sample (M=18.16 SD=3.331) place greater value on science than middle school students (M=17.27 SD=3.834), $t(370) = 2.25$, $p = 0.025$ (see table 9). Scores from the mATSI suggested middle school students experienced greater levels of anxiety toward science (M=12.44, SD=4.277) than high school students (M=11.06, SD=3.559), $t(370) = -3.16$, $p = 0.002$. And finally, high school students reported greater levels of science self-efficacy on the SLA (M=19.44, SD=3.097) than middle school students (M=19.32, SD=3.203), $t(370) = 2.59$, $p = 0.010$.

Table 6 *Description of Variables Examined for Each Construct*

Construct	**Instrument Source**	**Variables**
Demonstrated Science Literacy	SLA-D	Items 1 – 19
	Middle School General Science Instrument	Items 1 - 25
	High School General Science Instrument	Items - 17
Value of Science	SLA-MB	Items 1 - 6
	mATSI	Subscale Items 9 - 13
Self-Efficacy for Science Literacy	SLA-MB	Items 7 - 14
	mATSI	Subscale Items 14 -18
Source and Certainty of Scientific Knowledge	SLA-MB	Items 15 - 25
Perception of Science Teacher	mATSI	Subscale Items 1 - 3
Anxiety Toward Science	mATSI	Subscale Items 4 - 8
Interest in Science	mATSI	Subscale Items 19 –25
Interest in Science Career	STEM-CIS	Items 1 – 11

Table 7 *Descriptive Data for SLA-Motivation and Beliefs Items*

Survey Item	**Middle School Sample** N = 257				**High School Sample** N = 141			
	M	**SD**	**Skewness**	**Kurtosis**	**M**	**SD**	**Skewness**	**Kurtosis**
In general, I find working on science assignments.	3.19	1.019	-.211	-.198	3.40	1.034	-.310	-.369
Compared to most of your other activities, how useful is what you learn in science?	3.41	1.115	-.170	-.934	3.60	.992	-.290	-.754
For me, being good in science is (not at all important – Very important).	3.96	1.036	-.716	-.394	4.236	.8503	-.901	.399
Compared to most of your other activities, how important is it for you to be good at science?	3.48	1.022	-.206	-.605	3.754	.9473	-.555	-.069
How much do you like doing science?	3.35	1.222	-.311	-.896	3.468	1.1483	-.365	-.747
Some things that you learn in school help you do things better outside of class, that is, they are useful. For example, learning about plants might help you grow a garden. In general, how useful is what you learn in science?	3.56	1.088	-.441	-.514	3.574	1.0572	-.291	-.709
I know when to use science to answer questions.	3.58	.822	-.309	.426	3.70	.754	-.461	.638
I can use science to make decisions about my daily life.	2.99	1.046	-.026	-.496	3.113	.9717	-.136	-.158

Table 7 *Descriptive Data for SLA-Motivation and Beliefs Items (continued)*

Survey Item	**Middle School Sample** N = 257				**High School Sample** N = 141			
	M	**SD**	**Skewness**	**Kurtosis**	**M**	**SD**	**Skewness**	**Kurtosis**
I know how to use the scientific method to solve problems.	3.56	1.109	-.462	-.480	3.993	.8984	-.705	.402
It is easy for me to tell the difference between scientific findings and advertisements.	3.46	1.085	-.277	-.632	3.879	1.0452	-.517	-.777
When I do my work in science class, I am able to find the important ideas.	3.60	.999	-.422	-.240	3.652	.8700	-.513	.258
I can use math to answer scientific questions.	3.90	1.062	-.949	.497	4.113	.8871	-.661	-.461
I can tell the difference between observations and conclusions in a story.	4.10	.955	-1.041	.726	4.36	.795	-1.264	1.743
It is easy for me to make a graph of my data.	4.14	.970	-1.138	1.051	4.03	.902	-.886	.761
Everybody has to believe what scientists say.	1.90	1.001	.950	.224	1.76	1.027	1.182	.457
All questions in science have one right answer.	1.92	.977	1.044	.712	1.723	.8792	1.213	1.157
Scientific knowledge is always true.	2.09	.972	.632	-.204	2.099	1.0509	.623	-.503

Table 7 *Descriptive Data for SLA-Motivation and Beliefs Items (continued)*

	Middle School Sample N = 257				High School Sample N = 141			
Survey Item	**M**	**SD**	**Skewness**	**Kurtosis**	**M**	**SD**	**Skewness**	**Kurtosis**
In science, you have to believe what the science books say about stuff.	2.26	1.081	.580	-.381	2.014	1.0211	.625	-.620
The most important part of doing science is coming up with the right answer.	2.47	1.182	.373	-.729	2.376	1.1119	.286	-.910
Whatever the teacher says in science class is true.	2.33	1.090	.389	-.732	2.014	.9635	.651	-.113
Scientists pretty much know everything about science; there is not much more to know.	1.70	.961	1.466	1.752	1.27	.631	1.003	6.067
If you read something in a science book, you can be sure it's true.	2.39	1.081	.444	-.420	2.150	.8938	.428	-.233
Only scientists have a result from an experiment that is the only answer.	1.80	.881	1.063	.845	1.482	.7132	1.741	4.107
Scientists always agree about what is true in science.	1.91	1.023	1.028	.387	1.53	.841	1.723	2.739
Only scientists know for sure what is true in science.	1.74	.866	1.050	.543	1.50	.723	1.423	1.694

Table 8 *Mean and Standard Deviation of MATSI Items/Scales*

Scale	Middle School Sample (*N* = 237)		High School Sample (*N* = 135)	
Item	M	SD	M	SD
Perception of the teacher	**10.77**	**2.356**	**11.00**	**1.672**
Science teachers make science interesting.	3.38	1.009	3.52	.700
Science teachers present material in a clear and understandable way.	3.59	.932	3.46	.798
Science teachers are willing to give us individual help.	3.79	.928	4.03	.701
Anxiety toward science	**12.44**	**4.277**	**11.06**	**3.559**
When I hear the word "science," I have a feeling of dislike.	2.72	1.234	2.36	1.075
I feel tense when someone talks to me about science.	2.23	.995	1.96	.859
It makes me nervous to even think about doing science.	1.97	.976	1.76	.738
It scares me to have to take science class.	1.93	1.008	1.70	.802
I have a good feeling toward science.*	3.59	1.159	3.28	.920
Value of science to society	**17.27**	**3.834**	**18.16**	**3.311**
Science is useful for solving the problems of everyday life.	3.26	1.033	3.33	1.044
Most people should study some science.	3.37	1.036	3.69	.942
Science is helpful in understanding today's world.	3.78	.922	4.07	.878
Science is of great importance to a country's development.	3.78	.996	4.10	.863
It is important to know science in order to get a good job.	3.08	1.035	2.97	1.007
Self-confidence in science	**16.22**	**2.019**	**16.21**	**1.764**
I do not do very well in science.*	4.73	1.177	4.87	1.040
Science is easy for me.	3.42	1.138	3.39	1.051

I usually understand what we are talking about in science.	3.66	1.091	3.73	.981
No matter how hard I try, I cannot understand science.*	5.05	1.063	5.11	.975
I often thing, “I cannot do this,” when a science assignment seems hard.”	2.46	1.151	2.34	1.121
Desire to do science	**20.23**	**6.426**	**21.40**	**6.251**
Science is something which I enjoy very much.	3.08	1.236	3.33	1.106
I would like to do some reading in science which has not been assigned to me.	2.38	1.142	2.34	1.134
Sometimes I read ahead in our science book.	2.12	1.098	1.91	1.011
I like the challenge of science assignments.	2.74	1.241	2.92	1.191
It is important to me to understand the work I do in the science class.	3.93	1.012	4.13	.757
Science is one of my favorite subjects.	3.01	1.370	3.33	1.269
I have a real desire to learn science.	2.97	1.268	3.44	1.176

Table 9 *Descriptive Data for STEM-CIS*

	Middle School Sample N = 237				**High School Sample** N = 135			
Survey Item	**M**	**SD**	**Skewness**	**Kurtosis**	**M**	**SD**	**Skewness**	**Kurtosis**
I am able to get good grades in my science class.	4.05	.949	-.933	.591	4.15	.758	-.985	1.356
I am able to complete my science homework	4.14	.804	-.801	.353	4.27	.660	-.824	1.524
I plan to use science in my future career.	3.12	1.192	-.091	-.686	3.87	1.096	-.679	-.446
I will work hard in my science classes.	4.26	.723	-.772	.834	4.33	.611	-.532	.482
If I do well in science classes, it will help me in my future career.	3.62	1.109	-.437	-.503	4.01	1.008	-.825	.004
My parents would like it if I choose a science career.	3.00	1.064	.034	-.013	3.44	.952	.109	-.432
I am interested in careers that use science.	2.86	1.305	.114	-.971	3.70	1.185	-.634	-.524
I like my science class.	3.61	1.161	-.869	.049	3.76	1.009	-.702	.150
I have a role model in a science career.	2.55	1.172	.397	-.502	2.73	1.204	.328	-.633
I would feel comfortable talking to people who work in science careers.	3.14	1.153	-.200	-.592	3.60	1.073	-.429	-.410
I know of someone in my family who uses science in their career.	3.11	1.414	-.040	-1.297	3.66	1.288	-.615	-.786

Research Question 1

In order to investigate the degree to which the SLA is a valid measure of middle school students' general scientific knowledge, motivation, and beliefs related to science epistemology, exploratory factor analyses were conducted for the SLA-MB and 19-item version of the SLA-D administered to the middle school participants. Reliability analyses were then conducted by generating Cronbach's alpha for the SLA-MB overall as well as each subscale, and the SLA-D.

Exploratory Factor Analysis of the SLA-MB Administered to Middle School Students

A principle-axis factor analysis was conducted to determine the factor structure of the 25 item SLA-MB administered to middle school students within the study (see Table 6 for item descritions). Principal-axis extraction was utilized because factors were suspected to be more than minimally correlated (Crawford et al., 2010). First, mean imputation was utilized for missing data as recommend by Tabachnick and Fidell (2001). Normality was initially assessed by conducting a Shapiro-Wilk (1965) test, which indicated a non-normal distribution ($p < 0.001$). A visual examination was then conducted of the skewness and kurtosis of each item and the variance was found to be within acceptable guidelines of absolute skew values < 2 and absolute kurtosis values < 7 (West et al., 1996). Factor analysis is different from other multivariate procedures in that there is no identification of independent and dependent variables (Beavers et al., 2013). The procedure examines relationships between variables without consideration of one variables' influence on another. As a result, normality is not required when certain methods of extraction are utilized (Tabachnick & Fidel, 2001). The factorability of the items was tested through viewing the item correlation matrix, and by computing the Kaiser-Meyer-Olkin (KMO) index for the item set. A KMO index that is 0.6 or greater is considered to suggest good factorability (Tabachnick & Fiddell, 2001).

Literature addressing factor analysis contains a large amount of information regarding criteria for determining adequate sample size. These criteria vary greatly and are often debated (Beavers et al., 2013). While a minimum number of cases, or subject-to-variables ratio, has often been utilized in determining sample size requirements, the literature suggests that ratio criteria do not provide enough guidance (Osborne & Costello, 2004). Guadagnoli and Velicer (1988) suggest that adequate sample size is conditional upon the strength of the factors and the items. They propose sample size is not relevant if the factors have four or more items with loadings of 0.60 or higher. If the factors have 10 to 12 items with loadings of 0.40 or higher, then a sample size of 150 or more is needed. And finally, if factors have few variables with moderate to low loadings, a sample size of 300 is needed to interpret results with confidence. In following these recommendations, the adequacy of sample size was assessed by examining factor loadings after each factor analysis. Findings related to sample size are presented with the results of each factor analytic procedure.

An initial unrotated solution was conducted to estimate the number of factors suggested by the data. The Kaiser-Meyer-Olkin (KMO) measure verified the sampling adequacy for the analysis, KMO = 0.888 ('great' according to Hutcheson & Sofroniou, 1999). Bartlett's test of sphericity χ^2 (300) = 2693.84, $p < 0.001$, indicated the correlations between items were sufficiently large. Both parallel analysis and the scree plot showed inflexions that would justify a 3, 4, or 5 factor model (see Figure 1). Criteria for retaining the model with the best fit included: parallel analysis, a KMO score larger than 0.7, and at least 3 items with significant loadings, >.30 (Costello & Osborne, 2005).

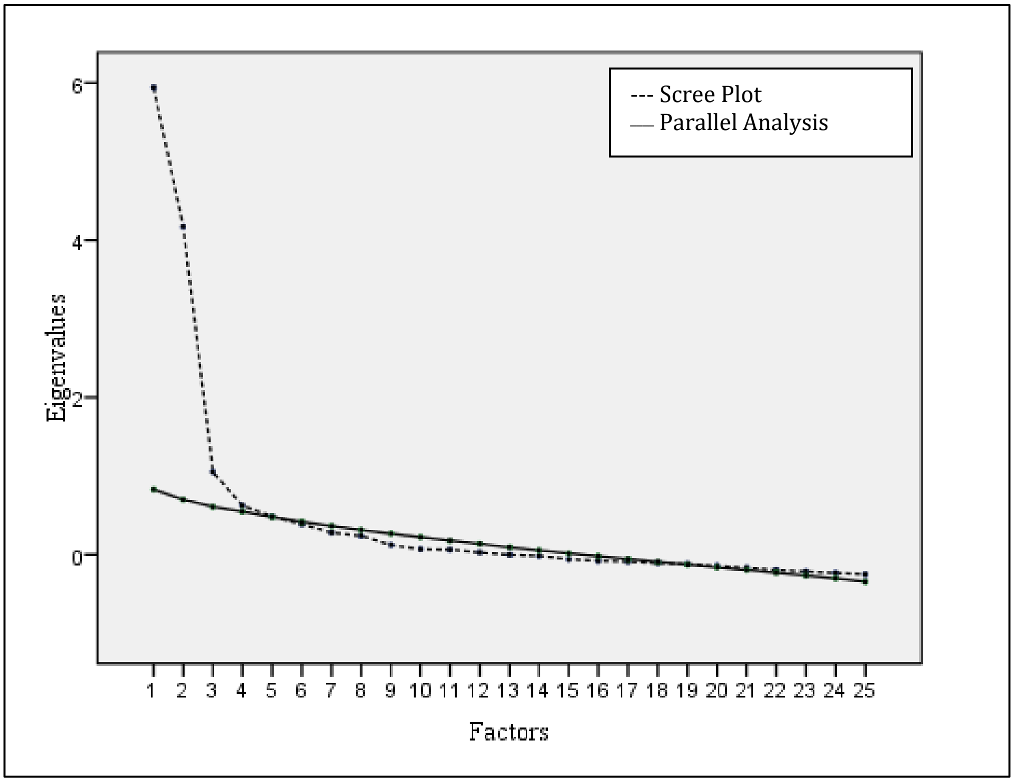


Figure 1 *Initial Unrotated Solution for SLA-MB Middle School Sample*

Oblique rotation was employed for each of the model assessments because it assumes the factors are correlated and produces a correlation matrix, which provides a way to confirm the correct rotation method was utilized (Costello & Osborne, 2005). The 5-factor model produced a KMO value of 0.888, explained 59.602% of the total variance, and the factor correlation matrix confirmed that oblimin rotation was correctly utilized (r = -.418 to 0.524). However, within this model there were only two items with significant loadings to the fifth factor.

The 4-factor model also produced a KMO value of 0.888. This model explained 55.615% of the total variance and the factor correlation matrix confirmed that oblimin rotation was correctly utilized s (r = -.274 to 0.521). The 4 factor model also failed to meet the criteria of containing three significant loadings, having only one item with a significant loading to the fourth component.

The 3-factor model was determined to provide the best fit with a KMO value of 0.888 and explaining 50.899% of the variance. There were no crossloadings within the model and each item significantly loaded to one of the three factors (see Figure 2). The factor correlation matrix confirmed that oblimin rotation was correctly utilized (r = 0.020 to 0.574). Factor one captured the most variance 25.834%, and contained the six items comprising the Value of Science scale plus an additional item "I can use science to make decisions about my daily life" from the Self Efficacy for Science Literacy scale. The second factor captured 18.799% of the total variance and contained the 11 items which comprise the Source and Certainty of Scientific Knowledge scale. The third factor captured the least amount of variance, 6.266%, and contained seven of the eight items from the Self Efficacy for Science Literacy scale (see Table 10). In following the guidance of Guadagnoli and Velicer (1988) on determining a minimum sample size, the 257 cases utilized for this analysis is deemed to be adequate.

Table 10 *Three Factor solution for the SLA-MB Middle School Sample*

Item	Loading		
	Factor 1	Factor 2	Factor 3
In general, I find working on science assignments (very boring – very interesting/fun).	.568	.204	.037
Compared to most of your other activities, how useful is what you learn in science?	.851	-.171	.066
For me, being good in science is (not at all important – Very important).	.530	.169	-.012
Compared to most of your other activities, how important is it for you to be good at science?	.717	-.025	-.019
How much do you like doing science?	.623	.169	.037
Some things that you learn in school help you do things better outside of class, that is, they are useful. For example, learning about plants might help you grow a garden. In general, how useful is what you learn in science?	.776	-.069	.071
I can use science to make decisions about my daily life.	.468	.217	-.100
I know when to use science to answer questions.	.154	.585	-.117
I know how to use the scientific method to solve problems.	.031	.659	-.021
It is easy for me to tell the difference between scientific findings and advertisements.	.114	.561	.005
When I do my work in science class, I am able to find the important ideas.	.042	.709	-.096
I can use math to answer scientific questions.	.049	.628	-.026
I can tell the difference between observations and conclusions in a story.	-.043	.657	.029
It is easy for me to make a graph of my data.	-.058	.580	.114
Everybody has to believe what scientists say.	.092	.239	.391
All questions in science have one right answer.	.068	.030	.527
Scientific knowledge is always true.	.132	.106	.577
In science, you have to believe what the science books say about stuff.	-.004	.064	.706
The most important part of doing science is coming up with the right answer.	-.046	.002	.630
Whatever the teacher says in science class is true.	.016	.088	.589
Scientists pretty much know everything about science; there is not much more to know.	-.048	-.105	.696
If you read something in a science book, you can be sure it's true.	.071	-.046	.682
Only scientists have a result from an experiment that is the only answer.	-.112	-.072	.695
Scientists always agree about what is true in science.	.021	-.154	.620
Only scientists know for sure what is true in science.	-.005	-.006	.702
Eigenvalues	6.456	1.567	4.700
Rotated Explained Variance	25.824	6.266	18.799

- Model explains 50.899% of variance
- $\alpha = .875$
  F1 $\alpha = .865$
  F2 $\alpha = .872$
  F3 $\alpha = .840$

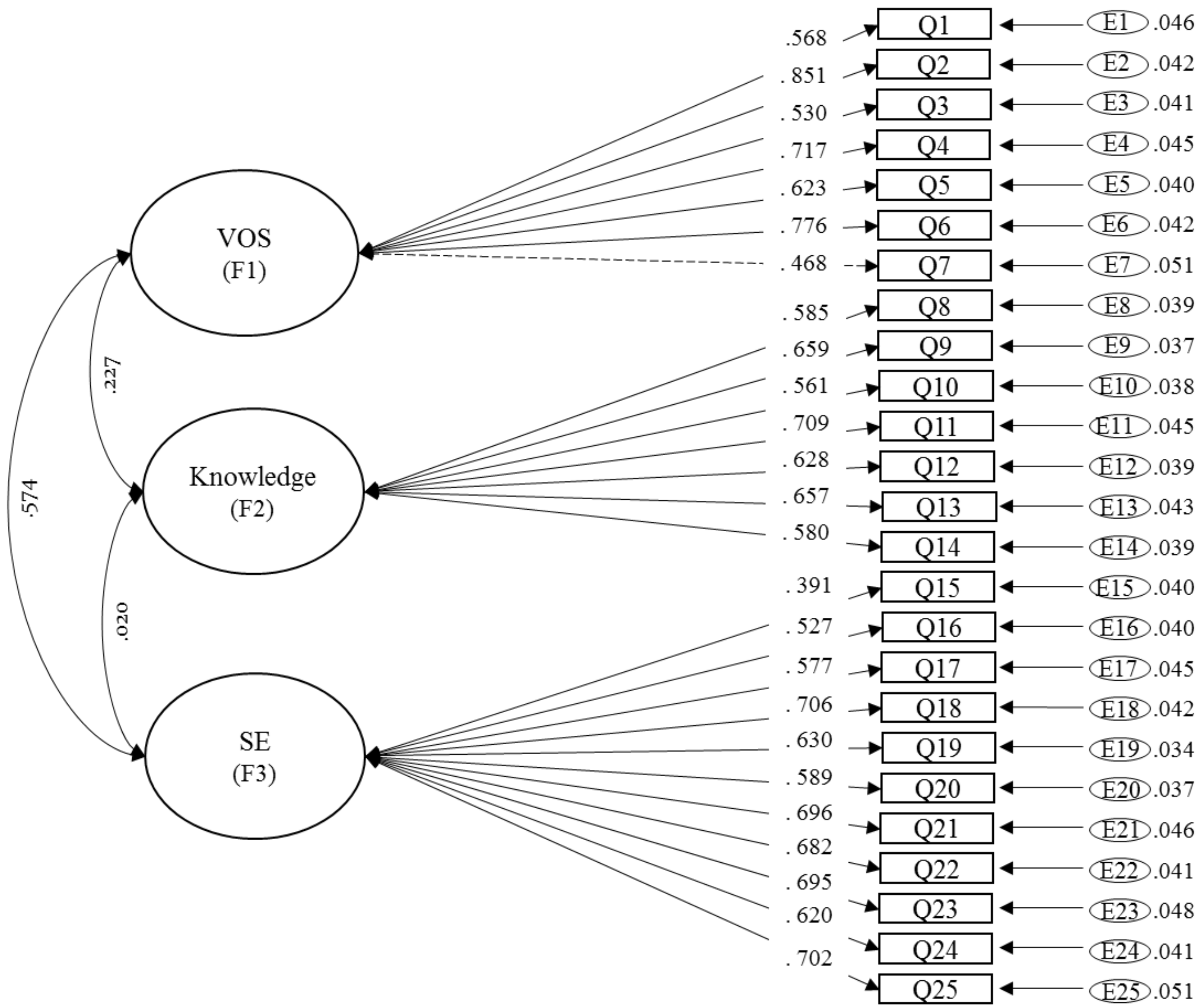


Figure 2 *Path Diagram of Three Factor Solution for the SLA-MB Middle School Sample*

Reliability – Item analysis. In assessing the reliability of the scales within the instrument, an assessment of variance was conducted by generating Cronbach's alpha. For exploratory research, an alpha value of 0.7 is generally considered 'adequate' with a value of 0.8 or higher considered as evidence of a 'good scale' (Dattalo, 2013). Assessment of internal consistency for all items within the instrument resulted in a coefficient of = 0.875. Examination of the individual factors also produced indicators of 'good' reliability; Factor 1 $\alpha = 0.865$, Factor 2 $\alpha = 0.872$, and Factor 3 $\alpha = 0.840$.

Exploratory Factor Analysis of the SLA-D Administered to Middle School Students

For the 249 responses from the middle school sample, the mean SLA-D score was 10 correct (53%) with a range of 2-19 (11-100%). A principle-component factor analysis was conducted to determine the factor structure of the 19 item SLA-D administered to middle school students within the study. Principal-component analysis was utilized for the procedure because Fives et al. (2014) had previously found the 25-item instrument to be unidimensional (Crawford et al., 2010). The Shapiro-Wilk test of normality indicated a non-normal distribution ($p < 0.001$). Examination of skewness and kurtosis of each item indicated the variance was within acceptable guidelines (West et al., 1996) with absolute skew values < 2 and absolute kurtosis values < 7. Prior to conducting the analysis, a correlation matrix was examined to evaluate the factorability of matricies. The majority of intercorrelatons fell below 0.30, suggesting that factoring may not be beneficial (Beavers et al., 2013). The initial extraction was conducted without rotation and produced a KMO score that fell below the previously determined threshold of 0.7 (.609). Additionally, parallel analysis provided supporting evidence that the instrument is unidimensional (see Figure 3). Because of these findings, further analysis utilizing rotation was not conducted.

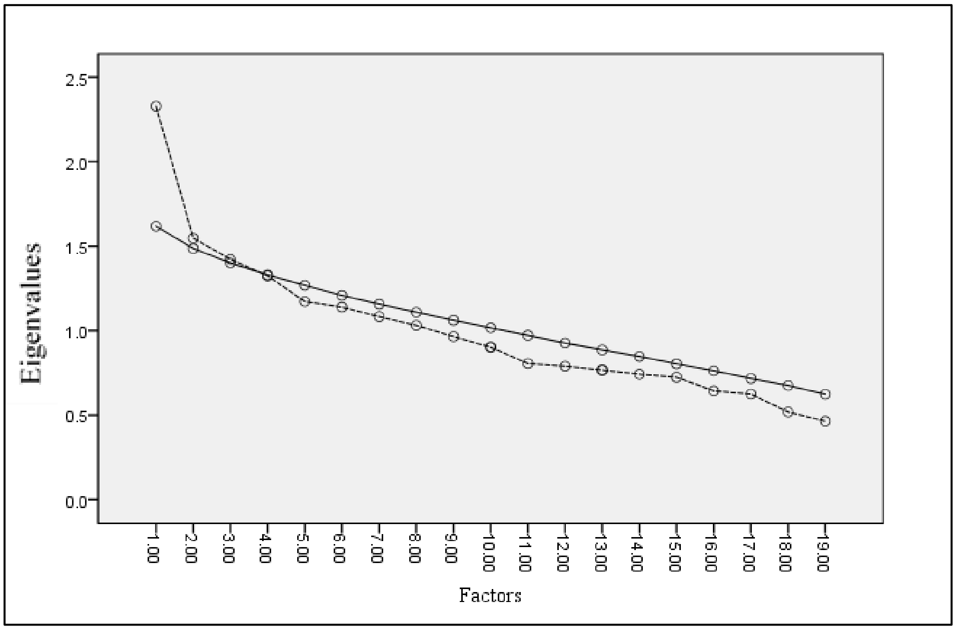


Figure 3 *Initial Unrotated Solution for SLA-D Middle School Sample*

Reliability – Item analysis. Considering the SLA-D as a single measure, assessment of internal consistency included all items within the instrument. The resulting alpha of 0.727 demonstrates evidence of 'adequate' reliability (Dattalo, 2013).

Research Question 2

To assess the degree to which the SLA is a valid measure of high school students' general scientific knowledge, motivation, and beliefs related to science epistemology, exploratory factor analyses were conducted for the SLA-MB and 19-item version of the SLA-D administered to the high school participants. Reliability analyses were then conducted by generating Cronbach's alpha for the SLA-MB overall as well as each subscale, and the SLA-D.

Exploratory Factor Analysis of the SLA-MB Administered to High School Students

A principle-axis factor analysis was again used to determine the factor structure of the 25 item SLA-MB administered to high school students within the study. The Shapiro-Wilk test of normality indicated a non-normal distribution ($p < 0.001$). Examination of skewness and kurtosis of each item indicated the variance was within acceptable guidelines (West et al., 1996)

with absolute skew values < 2 and absolute kurtosis values < 7 (see Table 7). The Kaiser-Meyer-Olkin (KMO) measure verified the sampling adequacy for the analysis, KMO = 0.784 ('middling' according to Hutcheson & Sofroniou, 1999). Bartlett's test of sphericity χ^2 (300) = 1398.28, $p < 0.001$, indicated the correlations between items were sufficiently large. Both the parallel analysis and the scree plot showed inflexions that would justify a 3 to 6 factor model (see Figure 4). Criteria for retaining the model with the best fit included: parallel analysis, a KMO score larger than 0.7, and at least 3 items with significant loadings, >.30 (Costello & Osborne, 2005).

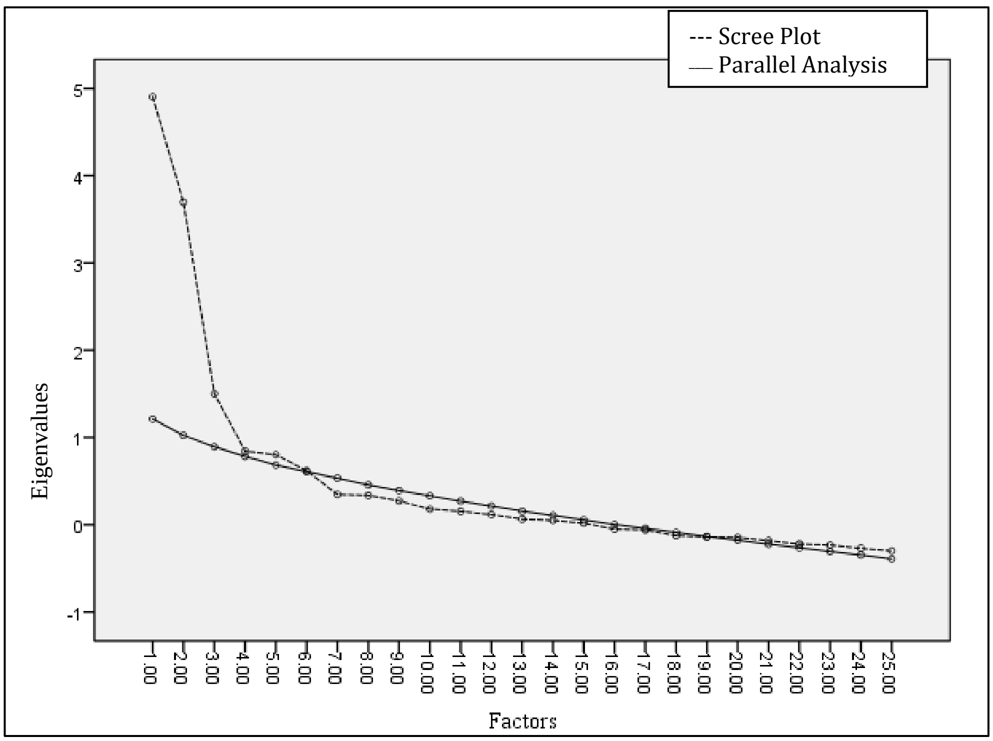


Figure 4 *Initial Unrotated Solution for SLA-MB High School Sample*

Oblique rotation was employed for each of the model assessments. The 6-factor model produced a KMO value of 0.784, explained 57.370% of the total variance, and the factor correlation matrix confirmed that oblimin rotation was correctly utilized (r = -.014 to 0.343).

Each of the six components had at least three significant loadings. Of the 25 items, seven had significant loadings to a second factor.

The 5- and 4- factor models also produced a KMO value of 0.784. The models explained 57.370% and 51.920% of the total variance respectively. The factor correlation matrix confirmed that oblimin rotation was correctly utilized (r = -.040 to 0.313). Of the 25 items, 3 items crossloaded to a second component within the 5-factor model, and 4-items crossloaded to a second component within the 4 factor model (see Figure 3).

The 3-factor model produced a KMO value of 0.784 and explained 46.422% of the variance. The factor correlation matrix confirmed that oblimin rotation was correctly utilized (*r* = 0.021 to 0.357). Factor one captured the most variance 21.369%, and contained the six items comprising the Value of Science scale plus an additional item "I can use science to make decisions about my daily life" from the Self Efficacy for Science Literacy scale (see Table 11). Two items from factor 3 crossloaded to Factor 1. The second factor captured 16.844% of the total variance and contained ten of the 11 items which comprise the SLA-MB Source and Certainty of Scientific Knowledge scale. The third factor captured the least amount of variance, 8.209% and contained seven of the eight items from the SLA-MB Self Efficacy for Science Literacy scale and one item from the SLA-MB Self Efficacy for Science scale. As with the middle school sample, the adequacy of the high school sample was assessed after the analysis utilizing criteria recommended by Guadagnoli and Velicer (1998). Factors 1 and 2 meet the threshold of producing four or more loadings of 0.60 or higher. However, Factor 3 falls short, producing no loadings of 0.60 or higher. Additionally, Factor 3 does not meet secondary threshold of 10 to 12 items that load moderately at 0.40 or higher. This would suggest an inadequate sample size for confidently interpreting the results of the factor analysis.

Table 11 *Three Factor solution for the SLA-MB High School Sample*

Item	Loading		
	Factor 1	Factor 2	Factor 3
In general, I find working on science assignments (very boring – very interesting/fun).	.788	-.039	.002
Compared to most of your other activities, how useful is what you learn in science?	.778	.016	.086
For me, being good in science is (not at all important – Very important).	.567	.018	-.046
Compared to most of your other activities, how important is it for you to be good at science?	.665	.166	-.085
How much do you like doing science?	.814	.001	-.022
Some things that you learn in school help you do things better outside of class, that is, they are useful. For example, learning about plants might help you grow a garden. In general, how useful is what you learn in science?	.698	-.129	.393
I can use science to make decisions about my daily life.	.333	-.068	.052
I know when to use science to answer questions.	.582	.036	.553
I know how to use the scientific method to solve problems.	-.002	-.045	.582
It is easy for me to tell the difference between scientific findings and advertisements.	.081	.032	.506
When I do my work in science class, I am able to find the important ideas.	.430	-.004	.524
I can use math to answer scientific questions.	-.034	-.034	.570
I can tell the difference between observations and conclusions in a story.	.010	.123	.514
It is easy for me to make a graph of my data.	-.050	.603	.104
Everybody has to believe what scientists say.	.102	.548	.216
All questions in science have one right answer.	-.070	.791	.006
Scientific knowledge is always true.	.219	.669	.043
In science, you have to believe what the science books say about stuff.	.057	.517	-.005
The most important part of doing science is coming up with the right answer.	-.159	.615	-.052
Whatever the teacher says in science class is true.	.032	.411	-.035
Scientists pretty much know everything about science; there is not much more to know.	-.046	.571	-.056
If you read something in a science book, you can be sure it's true.	.148	.533	-.008
Only scientists have a result from an experiment that is the only answer.	-.027	.242	-.309
Scientists always agree about what is true in science.	-.094	.451	-.262
Only scientists know for sure what is true in science.	-.016	-.039	.002
Eigenvalues	5.342	4.211	2.052
Rotated Explained Variance	21.369	16.844	8.209

- Model explains 46.422% of variance
- α = .808
  F1 α = .870
  F2 α = .828
  F3 α = .658

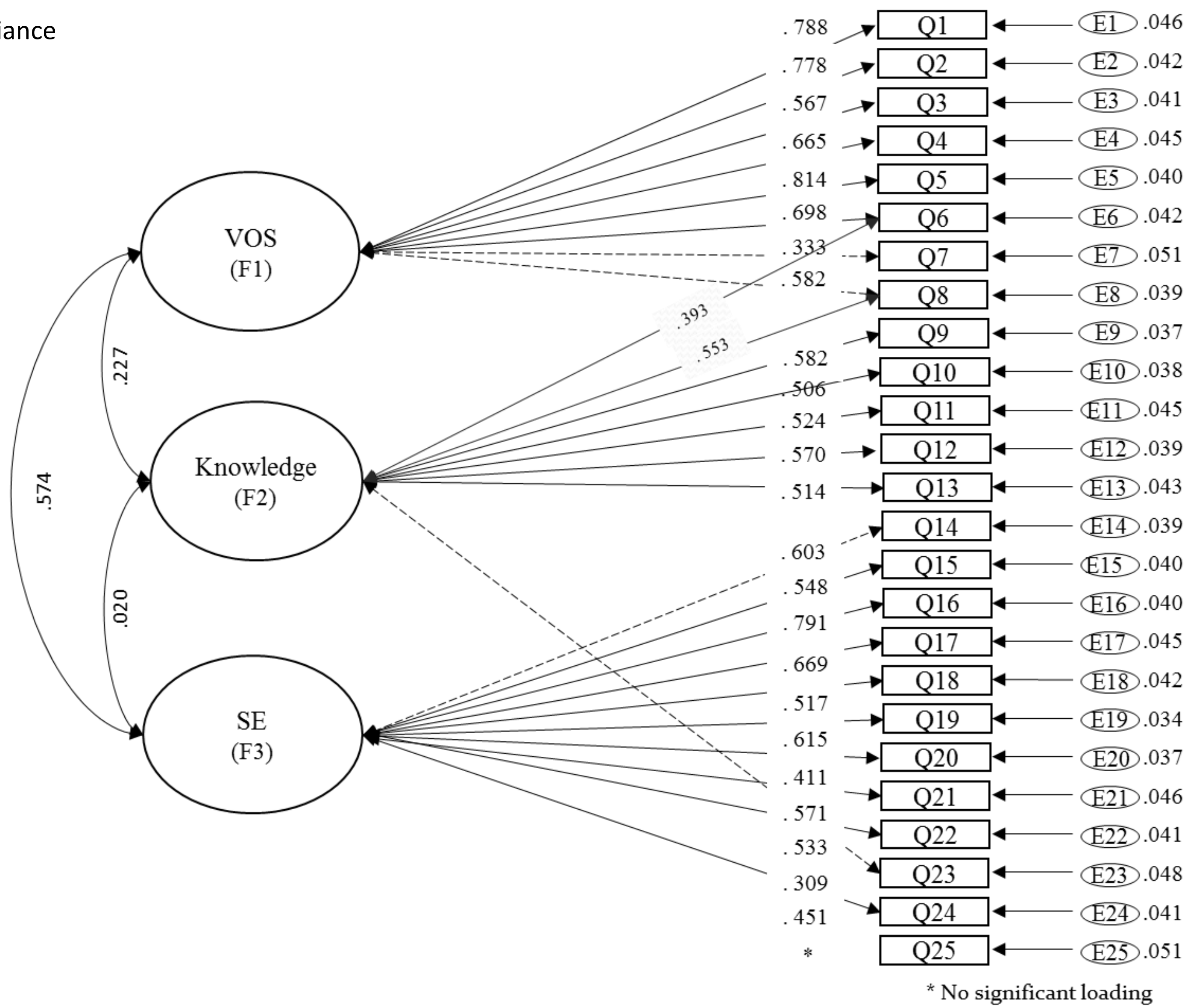


Figure 3 *Path Diagram of Three Factor Solution for the SLA-MB High School Sample*

Reliability – Item analysis. Assessment of internal consistency for all items within the instrument resulted in a Cronbach's alpha of 0.808. Examination of the individual components yielded 'good' reliability for Factor 1 ($\alpha = 0.870$) and Factor 2 ($\alpha = 0.828$), but Factor 3 produced a lower coefficient ($\alpha = 0.658$).

Exploratory Factor Analysis of the SLA-D Administered to High School Students

For the 142 responses from the high school sample, the mean SLA-D was 13 correct (68%) with a range of 3-19 (17-100%). A principle-component factor analysis was conducted to determine the factor structure of the 19 item SLA-D administered to high school students within the study. Principal-component analysis was utilized for the procedure because Fives et al. (2014) had previously found the instrument be unidimensional (Crawford et al., 2010). The Shapiro-Wilk test of normality indicated a non-normal distribution ($p < 0.001$). Examination of skewness and kurtosis of each item indicated the variance was within acceptable guidelines (West et al., 1996) with absolute skew values < 2 and absolute kurtosis values < 7. Prior to conducting the analysis, a correlation matrix was examined to evaluate the factorability of matricies. The majority of intercorrelatons fell below 0.30, suggesting that factoring may not be beneficial (Beavers et al., 2013). The initial extraction was conducted without rotation and produced a KMO score that fell below the previously determined threshold of 0.7 (.409). Because of these findings, further analysis utilizing rotation was not conducted.

Reliability – Item analysis. Considering the SLA-D as a single measure, assessment of internal consistency included all items within the instrument. The resulting alpha of 0.644 reaches the acceptable threshold for exploratory research (Dattalo, 2013).

Research Question 3

The extent to which scores on the SLA-MB correlate with scores on the Modified Attitudes Toward Science Inventory (mATSI) and the STEM Career Interest Survey (STEM-CIS) was investigated by generating a series of Pearson product-moment correlations. First, correlations were examined for all participants. Next, relationships between the SLA-MB, mATSI, and STEM-CIS were examined for the middle and high school samples separately.

Pearson Product-moment Correlations

Pearson product-moment correlations were run for all (N = 372) students in the study to determine the relationship between student overall scores on the SLA-MB and his/her overall scores on the mATSI and the STEM-CIS (see Table 11). There was a strong, positive correlation between SLA-MB and mATSI scores, which was statistically significant (r = 0.640, n = 372, p <.001). The relationship between scores on the SLA-MB and the STEM-CIS were also positive and statistically significant (r = 0.574, n = 372, p <.001). The relationship between scores on the SLA-MB and the STEM-CIS were also positive and statistically significant (r = 0.574, N = 372, p <.001), indicating students with higher scores on the SLA-MB expressed more interested in science careers. In reviewing the relationships between the three components of the SLA-MB with the STEM-CIS, two additional findings of interest emerged. The STEM-CIS was found to be positively correlated with the Value of Science scale (r = 0.724, N = 372, p <.001) and the Self Efficacy for Science scale (r = 0.640, N = 372, p <.001). These results suggest that students who place a higher value on science and students with greater self-efficacy for science are significantly more likely to be interested in science careers.

The analysis was then repeated separately for the middle and high school samples (see Table 9). A strong, positive, statistically significant correlation was found between middle

school student overall scores on the SLA-MB and MATSI (*r* = 0.654, *n* = 237, *p* <.001). The relationship between scores on the SLA-MB and the STEM-CIS for the middle school students were also positive and statistically significant (*r* = 0.603, *n* = 237, *p* <.001). The high school sample produced similar results, with strong positive correlations emerging between overall scores SLA-MB and MATSI scores (*r* = 0.603, *n* = 135, *p* <.001). The relationship between high school student overall scores on the SLA-MB and the STEM-CIS were also positive and statistically significant (*r* = 0.571, *n* = 135, *p* <.001). The SLA Value of Science scale indicated significant correlation with the mATSI Value of Science component (*r* = 0.627, N = 372, *p* <.001). The Self-Efficacy for Science component of the SLA was found to significantly correlate with the Self-Efficacy component of mATSI (*r* = 0.543, N = 372, *p* < 0.001).

Table 12 Correlations Between Student Scores on the SLA-MB, STEM-CIS, and mATSI

	Correlations		
	All Students (*N* = 372)	Middle School Students (*N* = 237)	High School Students (*N* = 135)
SLA-MB			
mATSI	.640*	.654*	.603*
STEM-CIS	.574*	.603*	.571*

* Correlation is significant at the 0.01 level (2-tailed).

Research Question 4

The extent to which student scores on the SLA-D predict scores on a student general science knowledge test was assessed by conducting bivariate linear regression for both the middle and high school samples. Prior to analysis, assumptions of linearity and normal distributions were checked and met. For both the SLA-D and middle school general knowledge instrument, a percent correct score was calculated for each student by dividing the number of questions answered by the number correctly answered. The general knowledge percent correct score was utilized as the dependent variable and the SLA-D percent correct score was the predictor variable.

Bivariate Linear Regression – Middle School Sample

Findings from the bivariate regression analysis for the middle school sample indicated that scores on the SLA-D (M = 0.543, SD = 0.199) were found to significantly predict scores on the general knowledge instrument (M = 0.627, SD = 0.159), F (1, 181) = 163.09 p < 001. Adjusted R^2 = 0.471. According to Cohen (1988) this is a medium effect size. The unstandardized regression weights indicate that when the percent of correct answers on the SLA-D increases by one unit, percent correct on the general knowledge instrument increases by 0.551.

Bivariate Linear Regression – High School Sample

As with the middle school sample, high school student's percent correct scores on the SLA-D (M = 0.690, SD = 0.158) were found to significantly predict scores on the general knowledge instrument (M = 0.612, SD = 0.161), F (1, 128) = 82.48 p < 001. Adjusted R^2 = 0.387. The unstandardized regression weights indicate that when the percent of correct answers on the SLA-D increases by one unit, percent correct on the general knowledge instrument increases by 0.614.

Summary

This chapter presented data analysis methods and findings for each research question. Exploratory factor analyses were conducted for the SLA-MB and SLA-D administered to the middle school sample. A 3-factor model provided the best fit for the SLA-MB with a KMO value of 0.888 and explaining 50.899% of the variance. Internal consistency indicators produced evidence of ‘good’ reliability; all items instrument resulted in a coefficient of $\alpha = 0.875$, Factor 1 $\alpha = 0.865$, Factor 2 $\alpha = 0.872$, and Factor 3 $\alpha = 0.840$. The SLA-D was found to be unidimensional, with the majority of intercorrelatons falling below 0.30, and a KMO score that fell below the previously determined threshold of 0.7 (.609).

Exploratory factor analyses were conducted for the SLA-MB and SLA-D administered to the high school sample. While a 3-factor model provided the best fit for the SLA-MB with a KMO value of 0.784, the model failed to produce evidence of an adequate sample size needed for confidently interpreting the results. The SLA-D was found to be unidimensional, with the majority of intercorrelatons falling below 0.30, and a KMO score that fell below the previously determined threshold of 0.7 (.409).

Pearson product-moment correlations indicated a strong, positive, statistically significant correlation between middle school student overall scores on the SLA-MB and MATSI ($r = 0.654$, $n = 237$, $p < .001$). The relationship between scores on the SLA-MB and the STEM-CIS for the middle school students were also positive and statistically significant ($r = 0.603$, $n = 237$, $p < .001$). The high school sample produced similar results, with strong positive correlations emerging between overall scores SLA-MB and MATSI scores ($r = 0.603$, $n = 135$, $p < .001$). The relationship between high school student overall scores on the SLA-MB and the STEM-CIS were also positive and statistically significant ($r = 0.571$, $n = 135$, $p < .001$). The SLA Value of

Science scale indicated significant correlation with the mATSI Value of Science component (r = 0.627, N = 372, p <.001). The Self-Efficacy for Science component of the SLA was found to significantly correlate with the Self-Efficacy component of mATSI (r = 0.543, N = 372, p < 0.001).

Findings from the bivariate regression analysis for the middle school sample indicated that scores on the SLA-D (M = 0.543, SD = 0.199) were found to significantly predict scores on the general knowledge instrument (M = 0.627, SD = 0.159), F (1, 181) = 163.09 p < 001. Adjusted R^2 = 0.471. As with the middle school sample, high school student's percent correct scores on the SLA-D (M = 0.690, SD = 0.158) were found to significantly predict scores on the general knowledge instrument (M = 0.612, SD = 0.161), F (1, 128) = 82.48 p < 001. Adjusted R^2 = 0.387.

Chapter 5

Discussion

The purpose of the current study was to examine validity evidence for the Science Literacy Assessment (Fives et al., 2014), a measure of student's ability to think scientifically as well as their motivation and beliefs related to science epistemology. The study extends previous validity evidence and furthers the initial work of Fives et al. (2014). Specifically, three sources of evidence were considered. First, the internal structure of each component of the SLA (demonstrated knowledge & motivation and beliefs) was examined using exploratory factor analysis. Next, concurrent criterion-related validity was examined using Pearson product-moment correlations to determine the relationship between student scores on the SLA-MB and his/her scores on the Modified Attitudes Toward Science Inventory, an instrument consisting of five scales: perceptions of the science teacher, anxiety toward science, value of science to society, self confidence in science, and desire to do science. Concurrent criterion-related validity was also examined using Pearson product-moment correlate to investigate the relationship between student scores on the SLA-MB and his/her scores on the STEM Career Interest Survey. And finally, evidence based on predictive criterion-related validity was explored for both samples (middle and high school) utilizing bivariate linear regressions to investigate the degree to which scores on the SLA-D predicted scores on a general science knowledge instrument.

Interpretation of Key Findings

Validity Evidence Based on Internal Structure

Internal structure evidence entails evaluating the relationship among individual assessment items and how they relate to the overarching construct as well as replicability, or reliability, across items (Cook et al., 2014). The factor structure of each of the two components

of the SLA was examined using exploratory factor analysis. By definition, a valid measure is also reliable (Dattalo, 2013), therefore internal consistency was examined for each of the two components through computation of Cronbach's alpha.

Motivation and beliefs: Middle school sample. Assessment of the underlying structure for the 25 item Motivation and Beliefs component of the SLA administered to the middle school sample revealed three distinct factors. This finding supports the initial conceptualization of the instrument as developed by Fives et al. (2014) measuring three, distinct components associated with scientific literacy. The Value of Science scale captured 25.834% of the variance, with the Source and Certainty of Science Knowledge scale explaining 18.799% of the variance, and the third factor, Self -Efficacy for Science Literacy, capturing 6.22%. In this study, 24 of the 25 items loaded to the same scale as when Fives et al. initially examined validity evidence for the instrument. The one item that loaded differently, "I can use science to make decisions about my daily life," was found by Fives et al. to be in the Value of Science scale, but was contained in the Self-Efficacy for Science scale in this study. It is important to note that factors 1 and 3 were found to be moderately correlated (r = 0.574) which could explain the loading discrepancy between studies. The method of extraction could also have caused the different loading. Fives et al. used principal components factor analysis, while this research utilized principal axis factoring.

In assessing the reliability of the scales, an assessment of variance was conducted by generating Cronbach's alpha. For exploratory research, an alpha value of 0.7 is generally considered 'adequate' with a value of 0.8 or higher considered as evidence of a 'good scale' (Dattalo, 2013). Assessment of internal consistency for all items within the instrument resulted in a coefficient of = 0.875. Examination of the individual factors also produced indicators of

‘good’ reliability; Value of Science α = 0.865, Self-Efficacy for Science α = 0.872, and Sources and Certainty of Scientific Knowledge α = 0.840.

Motivation and beliefs: High school sample. Analysis of the factor structure of the Motivation and Beliefs instrument administered to the sample of high school students was less conclusive than the middle school sample. The initial unrotated solution, parallel analysis, and examination of the scree plot warranted examination of models containing three to six factors. The 6-, 5-, and 4-factor models contained high numbers of cross-loadings (7, 3, & 4). Examination of the 3-factor solution had fewer cross-loadings (2), but only explained 46.422% of the total variance. The 3-factor solution also failed to produce evidence of an ample sample size, therefore the results should be interpreted cautiously. While the middle and high school samples could have been combined to address the low sample size, the research question driving the analysis focused on validity evidence surrounding the use of the instrument with high school students. Combining the samples would not have allowed for inferences about use with a high school population, and therefore, the samples were examined separately. Interestingly, the item “I can use science to make decisions about my daily life,” loaded to the Self-Efficacy for Science scale as it did with the middle school sample, instead of the Value of Science Scale as found by Fives et al.

Assessment of internal consistency for all items within the instrument resulted in a Cronbach’s alpha of 0.808. Examination of the individual components yielded ‘good’ reliability for Value of Science (α = 0.870) and Self-Efficacy for Science (α = 0.828), but Sources and Certainty of Scientific Knowledge produced a lower coefficient (α = 0.658).

Demonstrated knowledge: Middle school sample. During initial validation of the SLA, Fives et al. found the 25-item version of the demonstrated knowledge component to be

unidimensional. The current study sought to examine validity evidence for the 19-item version of the SLA-D. Principal-component factor analysis was conducted to examine the factor structure of the instrument administered to the middle school sample. An initial correlation matrix suggested factoring may not be beneficial as the majority of intercorrelations fell below 0.30 (Beavers et al., 2013). Initial unrotated extraction produced a KMO score (.609) that fell below the previously determined threshold of 0.7. Because these findings suggest a unidimensional scale, further analysis utilizing rotation was not conducted.

Assessment of internal consistency was conducted by considering the SLA-D as a single measure. The resulting alpha of 0.727 demonstrates evidence of 'adequate' reliability (Dattalo, 2013). The study findings suggest that, similar to the 25-item measure, the 19-item measure is also unidimensional. This could be advantageous to researchers who would like to utilize the instrument, but have concerns about the amount of time necessary for students to complete the 25-item version and the likelihood that it could be completed within one class period.

Demonstrated knowledge: High school sample. As with the middle school sample, the initial correlation matrix indicated factoring may not be appropriate as the majority of intercorrelations fell below.30 and the KMO (.409) index for this sample fell well below the threshold of 0.7. Additional analysis utilizing rotation was not conducted based on these findings.

Considering the SLA-D as a single measure, assessment of internal consistency included all items within the instrument. The resulting alpha of 0.644 reaches the acceptable threshold for exploratory research (Dattalo, 2013). Although findings suggest unidimensionality, and the reliability threshold was met, these findings should be considered in light of the probability that

the high school sample size is inadequate for making substantive conclusions based on the statistical procedures.

Concurrent Criterion-Related Validity Evidence

Relations with other variables is a source of validity evidence that establishes the statistical associations between assessment scores and another measure with a theoretical relationship (Cook et al., 2014). Concurrent-criterion related validity is one such source of evidence. A concurrent study obtains scores from the measurement tool and criterion scores at the same time (APA, 2013). In order to establish concurrent-criterion related validity evidence for the SLA, the extent to which scores on the SLA correlate with other established measures of student attitudes toward science, the mATSI and STEM-CIS, was examined. Findings indicate strong, positive correlations between the SLA-MB and mATSI scores ($r = 0.651$, N = 372, $p < .001$). Examination of corresponding scale items also produced strong correlations.

Middle school sample. Following examination of correlational evidence for all students in the study, the procedure was repeated for the middle school sample. A significant, positive relationship was found between overall SLA-MB and MATSI scores ($r = 0.654$, $n = 237$, $p < .001$). The SLA Value of Science scale indicated significant positive correlation with the mATSI Value of Science component ($r = 0.650$, N = 237, $p < .001$). And, the Self-Efficacy for Science component of the SLA was found to significantly correlate with the Self-Efficacy component of mATSI ($r = 0.543$, N = 237, $p < 0.001$).

The relationship between scores on the SLA-MB and the STEM-CIS were also positive and statistically significant ($r = 0.603$, $n = 237$, $p < 0.001$), indicating students with higher total scores on the SLA-MB expressed greater interested in science careers. The STEM-CIS was found to be positively correlated with the Value of Science scale ($r = 0.700$, N = 237, $p < 0.001$)

and the Self Efficacy for Science scale ($r = 0.659$, N = 237, $p < 0.001$). As with the findings for the total study sample, these results suggest that students who place a higher value on science and students with greater self-efficacy for science are significantly more likely to be interested in science careers.

High school sample. Strong positive correlations were found between SLA-MB and MATSI scores, which were statistically significant ($r = 0.641$, $n = 135$, $p < 0.001$). The SLA Value of Science scale indicated significant positive correlation with the mATSI Value of Science component ($r = 0.796$, N = 135, $p < 0.001$). And, the Self-Efficacy for Science component of the SLA was found to significantly correlate with the Self-Efficacy component of mATSI ($r = 0.560$, N = 135, $p < 0.001$).

The relationship between scores on the SLA-MB and the STEM-CIS were also positive and statistically significant ($r = 0.571$, $n = 135$, $p < 0.001$). The STEM-CIS was found to be positively correlated with the Value of Science scale ($r = 0.771$, N = 135, $p < 0.001$) and the Self Efficacy for Science scale ($r = 0.568$, N = 135, $p < 0.001$). Consistent with the above findings, these results suggest that students who place a higher value on science and students with greater self-efficacy for science are significantly more likely to be interested in science careers.

Predictive Criterion-Related Validity Evidence

The current study employed a second type of validity evidence based on relations with other variables, predictive criterion-related validity. A predictive study indicates the strength of the relationship between scores on the measurement tool and criterion scores obtained at a later time (APA, 2013). In order to examine the extent to which scores on the SLA-D predicted scores on a general science knowledge instrument, the SLA-D was administered during the first data

collection period and the general knowledge instrument was administered during the second data collection period.

Middle school sample. Bivariate linear regression was utilized to measure the relationship between middle school student's percent correct scores on the SLA-D and scores on a general knowledge instrument. Scores on the SLA-D (M = 0.627, SD = 0.199) were found to significantly predict scores on the general knowledge instrument (*M* = 0.627, *SD* = 0.159), F (1, 181) = 163.09 $p < 0.001$. Adjusted $R^2 = 0.471$, representing a medium effect size.

High school sample. Scores on the SLA-D (*M* = 0.690, *SD* = 0.158) were found to significantly predicted scores on the general knowledge instrument (*M* = 0.612, *SD* = 0.161), F (1, 128) = 82.48 $p < 0.001$. Adjusted $R^2 = 0.387$. While statistically significant, only approximately 39% of the variance of the SLA-D was found to be associated with the general science knowledge instrument.

Implications for Research and Practice

Literature regarding science literacy stresses the need to expand current methods for assessing the construct to include a more comprehensive perspective. This would allow for the development of instructional practices aligned to student interests and the ways they interact with science on a daily basis. However, the field of science education lacks instruments possessing the psychometric qualities required to serve that purpose (Osborne, Simon, & Collins, 2003). This issue is particularly pertinent in light of reports indicating that, when compared to countries around the world, adolescents in the United States rank below average in science literacy (OECD, 2014).

Supporting recommendations of more broadly defined assessments of science literacy are a variety of research studies which have linked student attitudes with learning, motivation, and

behavior. Given that classroom practices designed to promote affective and epistemological beliefs have been show to result in students' engaging more actively with science (Zepeda et al., 2015), it is important to recognize student attitudes as an important component of learning, Providing educators with an instrument that assesses students' motivation and beliefs toward science could serve as a valuable formative assessment tool in the classroom. From this perspective, instruction could be designed to impact student attitudes and beliefs about science in addition to addressing content.

The ability to assess science literacy within a broader context also introduces the potential of the SLA to be used by school counselors as well as classroom teachers to assess students' interests and aptitudes is necessary for providing students with guidance in developing their academic and career goals. The SLA could also be utilized by school counselors to assist students who are struggling in science classes by identifying areas which might benefit from academic intervention, such as self-efficacy for science or understanding science as a way of thinking.

Additional implications pertain to the use of the SLA as a classroom-based formative assessment tool. Teachers are encouraged to utilize data to inform practices and rely on assessment information about their individual students to guide their instructional process. Formative assessment, which provides frequent, interactive information about student progress, is essential to identifying learning needs and providing diagnostic information to teachers and students during the course of instruction. This information is also critical in assessing whether instructional practices are having the desired effect on student outcomes, allowing teachers to adapt teaching strategies to best serve the needs of their students. The SLA has the potential to serve this purpose by providing teachers with a formative assessment tool that assesses students'

understanding of science as an area of content as well as their understanding of science as an approach.

In addition to measuring scientific literacy within a broader context, a review of science education literature reveals the need to assess students' science knowledge that spans disciplines. While assessing content specific science knowledge is important, the argument can be made that measuring students' ability to understand science as an approach is more critical. Most students will not pursue a career in science as adults, however given our modern societal reliance on science and technology, they will interact with science to a great degree within the context of their daily lives. Providing educators with an instrument to measure students' ability to think scientifically would allow for educational opportunities that allow them to interact with scientific information needed to face practical issues and make informed decisions. The use of a tool such as the SLA could help facilitate the vertical alignment of curriculum, allowing educators to develop standards, assessments, and lessons that prepare students for success in the next grade level and beyond, and from one content area to the next.

Implications from this study are not limited to the classroom, but have the potential to shape assessment policies across the science education system. Assessment practices serve many diverse purposes, and no single assessment can meet the needs of all assessment goals. Use of the SLA allows for assessment policies and practices that integrate classroom-based formative assessment with standardized summative assessment to form a broader evaluation framework and/or a balanced assessment system. Such policy changes would involve not only educators in the classroom, but also state and school district administrators, those who make and implement policies at a local level.

The current study extends the validity evidence of the SLA and provides information with regard to student science learning and attitudes, the validity study findings also reveal several practical implications. Following the development of the SLA-D, Fives et al. expressed concerns about the length of the 25-item component measuring demonstrated science literacy and the likelihood of students' ability to complete the measure in one class session. This dissertation research contributes to the field of science education by providing evidence that the shorter, 19 item version is a valid instrument for measuring the science literacy of middle school students.

Study Limitations

As with all research, findings from this dissertation must be considered in light of the study's limitations. This investigation has several limitations that should be acknowledged. First, data were obtained from a nonprobability, convenience sample. Nonprobability samples introduce bias that could result in a sample that poorly represents the population. Convenience samples can lead to the under- or over-representation of particular groups within the sample. This weakness limits the generalizability of findings and constrains the ability to make valid inferences about the larger population (Agresti & Finlay, 2009).

Second, this study utilized self-report measures. One of the most significant limitations inherent to such survey measures is social desirability bias (Dillman, Smyth, & Christian, 2009). Social desirability bias occurs when participants respond in a manner that is perceived as socially acceptable, even if the responses are not reflective of their true feelings. In this study, participant responses may not have accurately represented their motivation and beliefs related to science. The methodology utilized for the current study attempted to mitigate social desirability bias by

ensuring data collection was conducted anonymously with no mechanism for connecting responses to student identities.

Third, while there is much debate within the literature about what constitutes adequate sample size for factor analysis, the high school sample utilized within this study failed to produce evidence of a sufficient number of observations required to produce a stable solution for the Motivation and Beliefs scale (Guadagnoli & Velicer, 1988). Therefore, valid inferences cannot be made about the factor structure of the SLA-MB administered to the high school students in this study. For the current research, the middle and high school samples were not combined in an attempt to address the small size of the high school sample because the research question focused on examining validity evidence supporting the use of the instrument with high school students.

Finally, the teacher and student participants may not be representative of the larger population. Teachers attending the CRESST Academy faced barriers which may have prevented participation in the study. These barriers include loss of instructional time, division level policies prohibiting student participation in research, and logistical or scheduling conflicts. Additionally, high school participants included students within a Health Science Chemistry class that is offered as part of a Health Science Specialty program. The program is geared toward students who are interested in pursuing careers within the healthcare industry. Given the possibility that the participants may not be representative of the larger population, generalizability of findings from this study are limited.

Recommendations for Future Research

The present study addresses a gap in the literature by contributing to the validity evidence of Science Literacy Assessment. Future research should seek to continue validation of the

instrument. Such research should aim to extend these findings to other samples, such as high school and post-secondary students. The participants for this investigation were limited to five schools in Virginia, limiting generalizability of the findings. Validity evidence should be expanded by administering the SLA to students in other states and outside of the United States in order to support widespread use of the instrument.

Additional forms of evidence could also be explored. For example, evidence based on test consequences was not examined within this study or the original validity evidence established by Fives et al. This study examined evidence based on relations to other variables by comparing scores on the SLA to instruments with similar theoretical connections. Evidence from sources more reflective of student learning, such as improvement in grades or measures of engagement, could also prove to be beneficial.

The various ways in which the instrument is utilized might also serve as an area of investigation. The field of science education could benefit from an investigation of the usefulness of the instrument for teachers in informing classroom practices. For example, the SLA should be evaluated as a pre-post assessment. This would facilitate the evaluation of various instructional strategies, such as those designed to increase students' science self-efficacy or ability to think about science in a way that span disciplines.

Conclusion

Findings from this dissertation suggest the SLA has appropriate psychometric properties for use in assessing middle school students' ability to think scientifically as well as their motivation and beliefs related to science. Exploratory factor analysis supported the underlying factor structure of the Motivation and Beliefs component as conceptualized by Fives et al. and measuring three distinct factors associated with science literacy: value of science, sources and

certainty of scientific knowledge, and self-efficacy for science. Results of this study replicated previous findings supporting the unidimensionality of the Demonstrated Knowledge component of the SLA. On the basis of validity analysis, reliability coefficients, and correlation studies, it appears that the SLA demonstrates good validity and reliability for use with middle school students. Results from the exploratory factor analysis on the Motivation and Belief scale for the high school sample cannot be interpreted with confidence because of the small sample size. Further research is needed to determine if the scale is appropriate for use with high school students.

The findings from this study are anticipated to help educators seeking to measure students' scientific literacy from a broader perspective. By aligning classroom instruction with the interests of students, and making science relevant to their every-day lives, teachers may be able to more fully engage students in science education and positively affect important attitudes, such as self-efficacy for science. Most students in the United States interact daily with science and technology in ways they may not even recognize. Equipping them to navigate these interactions in informed and purposeful ways, and to meaningfully participate in decisions that affect society at large would be noble accomplishment for science education.

References

Agresti, A., & Finlay, B. (2009). *Statistical Methods for the Social Sciences (4th ed.)*. Upper Saddle River, NJ: Prentice Hall.

Airasian, P.W. (1987). State mandated testing and educational reform: Context and consequences. *American Journal of Education, 95*(3), 393-314.

Aldridge, B.G. (1992). Project on scope, sequence, and coordination: A new synthesis for improving science education. *Journal of Science Education and Technology*, *1*(1), 13-21.

American Association for the Advancement of Science. (1989). *Science for All Americans*. Washington, D.C., Oxford University Press.

American Association for the Advancement of Science. (1992). Update Project 2061: Education for a changing future. Washington, DC.

American Educational Research Association. (2006). Standards for reporting on empirical social science research in AERA publications. *Educational Researcher*, *35*(6), 33-40.

American Psychological Association. (2013). *Publication manual of the American Psychological Association* (6th ed.). Washington, D.C.: American Psychological Association.

Anderson, K. (2012). Science education and test-based accountability: Reviewing their relationship and exploring implications for future policy. *Science Education*, *96*(1), 104-129.

Anelli, C. (2011). Scientific literacy: What is it, are we teaching it, and does it matter? *American Entomologist*, 57(4), 235-243.

Aronson, I. (2007). Negotiating the terrain of high-stakes accountability in science teaching (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Database. (UMI No. 3297419)

Atkinson, J. W. (1957). Motivational determinants of risk taking behavior. *Psychological Review* 64: 359-372.

Baker, V. (2004). Beyond appearances: Students' misconceptions about basic chemical ideas (2nd ed.). London: Royal Society of Chemistry.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Beavers, A.S., Lounsbury, J.W., Richards, J.K., Huck, S.W., Skolits, G.J., & Esquivel, S.L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research & Evaluation, 18*(6), 1-13.

Blalock, C. L., Lichtenstein, M. J., Owen, S., Pruski, L., Marshall, C., & Toepperwein, M. (2008). In pursuit of validity: A comprehensive review of science attitude instruments 1935–2005. *International Journal of Science Education*, *30*(7), 961-977.

Bloom, B.S. (1956). *Taxonomy of Educational Objectives Book*. Boston, MA: Allyn & Bacon.

Bloom, Benjamin S., Allison Davis, and Robert Hess (1965). *Compensatory education for cultural deprivation*. New York: Holt, Rinehart & Winston.

Britner, S., & Pajares, F. (2001). Self-efficacy beliefs, motivation, race, and gender in middle school science. *Journal of Women and minorities in Science and Engineering, 7*, 296-283.

Britner, S., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching*, *43*(5), 485-499.

Brooks, H. (1991). Scientific literacy and the future labour force. In T. Husen & J. P. Keeves (Eds.), Issues in science education (pp. 19–32). Oxford: Pergamon Press.

Brookhart, S.M. (2013). The public understanding of assessment in educational reform in the

United States. *Oxford Review of Education*, *39*(1), 52-71.

Bybee, R. W. (1997). Towards an understanding of scientific literacy. *Scientific literacy*, 37-68.

Bybee, R., McCrae, B., Laurie, R. (2009). PISA 2006: An assessment of scientific literacy. *Journal of Research in Science Teaching, 46*(8), 865-883.

Chen, J., & Pajares, F. (2010). Implicit theories of ability of Grade 6 science students: Relation to epistemological beliefs and academic motivation and achievement in science. *Contemporary Education al Psychology, 35*, 75-87.

Coble, J. (2006). Curricular constraints, high-stakes testing and the reality of reform in high school science classrooms (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Database. (UMI No. 3207430)

Collins, A. (1998). National science education standards: A political document. *Journal of Research in Science Teaching, 35* (7), 711-727.

Cohen, J. (1988). *Statistical power and analysis for the behavioral sciences* (2nd ed.). New York, NY: Rouledge/Taylor & Francis.

Cook, D.A., Zendejas, B., Hamstra, S.J., Hatala, R., & Brydges, R. (2014). What counts as validity evidence? Examples and prevalence in a systematic review of simulation-based assessment. *Advances in Health Science Education, 19*, 233-250.

Costello, A.B., & Osborne, J.W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation, 10*, 1-9.

Crawford, A.V., Green, S.B., Levy, R., Lo, W., Scott, L., Svetina, D., & Thompson, M.S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement, 70*(6), 885-901.

Cresswell, J.W. (2013). *Research Design: Qualitative, quantitative, and mixed methods approaches* (4th ed.) Thousand Oaks, CA.Sage Publications, Inc.

Cunningham, C. J., Weathington, B. L., & Pittenger, D. J. (2013). *Understanding and conducting research in the health sciences*. John Wiley & Sons.

Dattalo, P. (2013). *Analysis of multiple dependent variables*. New York, NY: Oxford University Press.

DeBoer, G.E. (1991). *A history of ideas in science education: Implications for practice*. New York: Teachers College Press.

DeBoer, G.E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching*, *37*(6), 582-601.

DeNovellis, R. L., & Lewis, A. J. (1974). *Schools become accountable: A PACT approach*. Washington, DC: Association for Supervision and Curriculum Development.

Dewey, J. (1902). *The Child and the curriculum*. Chicago, IL: The University of Chicago Press.

Dewey, J. (1910). Science as subject-matter and as method. *Science, 31*(787), 121-127.

Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. New York, NY: Simon & Schuster Inc.

Dewey, J. (1934). The supreme intellectual obligation. *Science Education, 18*(1), 1-4.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method (3rd ed.)*. New York, NY: John Wiley & Sons

Eccles, J.S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*, 109-132.

Eisenhart, M., Finkel, E., & Marion, S.F. (1996). Creating the conditions for scientific literacy: A

re-examination. *American Educational Research Journal*, *33*, 261-295.

Elder, A.D. (2002). Characterizing fifth grade students' epistemological beliefs in science. In B.K. Hofer & P.R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*, Mahwah, NJ: Lawrence Erlbaum, (347–364).

Enders, C.K. (2010). *Applied Missing Data Analysis*. New York: The Guilford Press.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C. & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272-299.

Falk, J.H. (2001). *Free-choice science education: How we learn science outside school*. New York: Teachers College Press.

Feinstein, N. (2011). Salvaging science literacy. *Science Education*, *95*(1), 168-185.

Feinstein, N. (2014). Making sense of autism: Progressive engagement with science among parents of young, recently diagnosed autistic children. *Public Understanding of Science*, *23*(5), 592-609.

Feinstein, N., Allen, S., & Jenkins, E. (2013). Outside the pipeline: Reimagining science education for nonscientists. *Science*, *340*, 314-317.

Fives, H., Huebner, W., Birnbaum, A. S., & Nicolich, M. (2014). Developing a measure of scientific literacy for middle school students. *Science Education*, *98*(4), 549-580.

Font-Rivera, M. J. (2003). *A descriptive study of the reported effects of state-mandated testing on the instructional practices and beliefs of middle school science teac*hers (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Database. (UMI No. 3091147).

Gabel, L. (1976). *The development of a model to determine perceptions of scientific literacy*.

(Doctoral dissertation). The Ohio State University, Columbus, Ohio.

Galton, M. (2002). Continuity and progression in science teaching at key stages 2 and 3. *Cambridge Journal of Education*, 32(2), 249 – 265.

Goetz Shuler, S., Backman, J., & Olson, S. (2009). The role of assessments and accountability. In B. B. Berns & J. O. Sandler (Eds.), Making science curriculum matter: Wisdom for the reform road ahead (pp. 49 – 59). Thousand Oaks, CA: Corwin Press.

Graber, S. (2007). *Sputnik and the Dawn of the Space Age*. NASA. Retrieved from http://history.nasa.gov/sputnik/ January 4, 2015.

Great Schools Partnership. In S. Abbott (Ed.), *The glossary of education reform*. Retrieved August 2014 from http://edglossary.org/hidden-curriculum.

Guadagnoli, E.& Velicer, W.F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103*, 265-275.

Hoffer, B. & Pintrich, P. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88-140.

Hoffer, B. & Pintrich, P.(Eds.) check APA (2002). *Personal Epistemology: The psychology of beliefs about knowledge and knowing.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hofstein, A., Eilks, I., & Bybee, R. (2011). Societal issues and their importance for contemporary science education –A pedagogical justification and the state-of-the-art in Israel, Germany, and the USA. *International Journal of Science and Mathematics Education, 9*, 1459-1483.

Holbrook, J., & Rannikmae, M. (2007). The nature of science education for enhancing scientific literacy. *International Journal of Science Education*, *29*(11), 1347-1362.

Holbrook, J., & Rannikmae, M. (2009). The meaning of science literacy. *International Journal of Environmental & Science Education*, *4*(3), 275-288.

Hurd, P. (1958). Science literacy: It's meaning for American schools. *Educational Leadership, 16*, 13-16, 52.

Hutcheson, G., & Sofroniou, N. (1999). *The multivariate social scientist*. London: Sage.

Impey, C., Buxner, S., Antonellis, J., & King, C. (2011). A twenty-year survey of science literacy among college undergraduates. *Journal of College Science Teaching*, 40(4), 31-37.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, 73(2), 509-527.

Jenkins, E. W. (2003). Environmental education and the public understanding of science. *Frontiers in Ecology and the Environment*, *1*(8), 437-443.

Joo, Y. J., Bong, M., & Choi, H. J. (2000). Self-efficacy for self-regulated learning, academic self-efficacy, and Internet self-efficacy in Web-based instruction. *Educational Technology Research and Development*, 48, 5-17.

Kahraman, N., & Sungur-Vural, S. (2014). The contribution of gender, socio-economic status and socio-cultural influence to Turkish students' task value beliefs in science. *Research in Education*, 91(1), 30-44.

Katzmann, J. M. (2007). *The influences of implementing state-mandated science assessment on teacher practice* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Database. (UMI No. 3280260)

Kersaint, G., Borman, K. M., Lee, R., & Boydston, T. L. (2001). Balancing the contradictions

between accountability and systemic reform. *Journal of School Leadership,* 11(3), 217 – 240.

Koballa, T., & Glynn, S. (2007). Attitudinal and motivational constructs in science learning. In S. Abell & N. Lederman (Eds.), *Handbook of research on science education*. Mahwah, NJ: Lawrence Erlbaum Associates.

Krajcik, J. Codere, S., Dahsah, C., Bayer., & Mun, K. (2014). Planning instruction to meet the intent of the Next Generation Science Standards. *Journal of Science Teacher Education*, *25*, 157-175.

Laetsch, W.M. (1987). A basis for better public understanding of science. In GIBA Foundational Conference (Ed.), Communicating science to the public. New York: John Wiley & Sons (1-18).

Laugksch, R.C. (2000). Scientific literacy: A conceptual overview. *Science Education*, *84*, 71-94.

Layton, D., Davey, A., & Jenkins, E. (1986). Science for special social purposes (SSSP): Perspectives on adult scientific literacy. *Studies in Science Education*, *13*, 27-52.

Lederman, N.G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching*, *29*(4), 331-359.

Liu, X. (2009). Beyond Science Literacy: Science and the public. *International Journal of Environmental & Science Education*, *4*(3), 301-311.

Marshall, J.C., & Alston, D.M. (2014). Effective, sustained inquiry-based instruction promotes higher science proficiency among all groups: A 5-year analysis. *Journal of Science Teacher Education*. *25*, 807-821.

Martin, L.M.W. (2004). An emerging research framework for studying informal learning and

schools. *Science Education, 88*(suppl. 1), S71-S82.

Miller. J. D , (1983). Scientific literacy: a conceptual and empirical review. Daedalus, 112(2), 29-48.

Miller, J. (1992). Toward a scientific understanding of the public understanding of science and technology. *Public* Understanding of Science, 1, 23-26.

National Academy of Sciences, National Academy of Engineering & Institute of Medicine. (2007). *Rising above the gathering storm: energizing and employing America for a brighter economic future*, National Academies Press.

National Aeronautics and Space Administration, Apollo 11 Mission Report retrieved from http://www.history.nasa.gov/alsj/a11/A11_PAOMissionReport.html January 4, 2015.

National Center for Educational Statistics (2012). *The Nation's Report Card: Science 2011* (NCES 2012-465). Washington, DC: U.S. Department of Education.

NGSS Lead States (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.

National Research Council. (1994, November). *National science education standards: Draft*. Washington, DC: National Academy Press.

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

National Research Council, (2001). Inquiry and the national science education standards. Washington, DC: National Academies Press.

National Research Council. (2002). *Scientific research in education.* Committee on Scientific Principles for Education Research. Shavelson, R.J., and Towne, L., Editors. Center for

Education. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

National Research Council. (2007). Taking science to school: Learning and teaching science in Grades K-8. Washington, DC: The National Academic Press. National Research Council (2012), *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press, Washington, DC.

National Research Council. (2011). *A Framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council. (2013). *The Next Generation Science Standards: Executive summary*. NGSS Lead States. (2013). Next Generation Science Standards: For states, by states. Washington, DC: The National Academies Press.

National Science Board (2006). *America's pressing challenge – Building a stronger foundation: A companion to Science and Engineering Indicators 2006*. Washington, DC: Government Printing Office.

National Science Board (2010). *Preparing the next generation of STEM innovators: Identifying and developing our nation's human capital*. Washington, DC: Government Printing Office.

National Science Board (2016). *Science and Engineering Indicators 2016*. Arlington, VA: National Science Foundation (NSB-2016-1).

National Science Teachers Association (1971). NSTA Position statement on school science education for the 70's. *The Science Teacher*, *38*, 46-51.

National Science Teachers Association. (1992). Scope, sequence, and coordination of secondary school science (Vol. I). Washington, DC.

National Science Teachers Association. (1995). Scope, sequence, and coordination of secondary school science (Vol. III). Arlington, VA: Author.

National Science Teachers Association (2010). *Position statement*. Washington DC: National Science Teachers Association.

Organization for Economic Co-Operation and Development (2014). PISA 2012 Results: *What students know and can do – Student performance in mathematics, reading and science* (Volume 1, Revised edition, February 2014), PISA, OECD Publishing.

Pella, M.O., O'Hearn, G.T., & Gale, C.W. (1966). Referents to scientific literacy. *Journal of Research in Science Teaching*, *4*, 199-208.

Penfield, R. D., & Lee, O. (2010). Test-based accountability: Potential benefits and pitfalls of science assessment with student diversity. Journal of Research in Science Teaching, 47(1), 6 – 24.

Petty, R. E., & Cacioppo, J. T. (1996). *Attitudes and persuasion: Classic and contemporary approaches*. Westview Press.

Pinder, P. J. (2008). A critique analysis of NCLB, increased testing, and past Maryland mathematics and science HAS exams: What are Maryland's practitioner's perspectives? Paper presented at the 16th annual Association for Science Teacher Education Conference, St. Louis, MO.

Pintrich, P.R., & Schunk, D.H. (2002). *Motivation in education: Theory, research and applications*. NJ: Merrill Prentice Hall.

Ricco, R., Pierce, S.S., & Medinilla, C. (2010). Epistemic Beliefs and achievement Motivation

in Early Adolescence. *Journal of Early Adolescence, 20*(2), 305-340.

Reynolds, A.J., (1991). The middle schooling process: Influences on science and mathematics achievement from the longitudinal study of American youth. *Adolescence, 26*, 132-157.

Reynolds, A.J., & Walberg, H.J. (1992). A structural model of science achievement and attitude: An extension to high school. *Journal of Educational Psychology, 84*(3), 371-382.

Roberts, D. (2007). Scientific literacy/science literacy. In Abel, S.K. & Lederman, N.G. (eds), *Handbook of research on science education*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Roth, W-M. & Lee, S. (2002). Scientific literacy as collective praxis. *Public Understanding of Science, 11*, 33-56.

Roth, W-M. & Lee, S. (2004). Science education as/for participation in the community. *Science Education*, *88*, 263-291.

Rudolph, J. (2014). Dewey's "Science as Method" a century later: Reviving science education for civic ends. *American Educational Research Journal, 51*(6), 1056-1083.

Shamos, M. (1995). *The myth of scientific literacy*. New Brunswick, NJ: Rutgers University Press.

Shapiro, S. S., & Wilk, M.B. (1965). *An analysis of variance test for normality (complete samples)*. Biometrika, 52, 591-611.

Shen, B. (1975a). Science literacy and the public understanding of science. In S. B. Day (Ed.), *Communication of scientific information*. Basel, Switzerland: S. Karger A.G. (44-52).

Shen, B. (1975b). Science literacy. *American Scientist*, *63*, 265-268.

Schommer, M. (1993). Epistemological development and academic performance among secondary students. *Journal of Educational Psychology, 85*(3), 406-411.

Shortland, M. (1988). Advocating science: Literacy and public understanding. *Impact of Science on Society*, *38*(4), 305-316.

Sinatra, G. M., Kienhues, D., & Hofer, B. K. (2014). Addressing challenges to public understanding of science: Epistemic cognition, motivated reasoning, and conceptual change. *Educational Psychologist*, *49*(2), 123-138.

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *Journal of Educational Research, 95*(6), 323-332.

Sirotnik, K. A. (2004). *Holding accountability accountable: What ought to matter in public education.* New York: Teachers College Press.

Smith, C., Maclin, D., Houghton, C., & Hennessey, M. (2000). Sixth-grade students' epistemologies of science: The impact of school science experiences on epistemological development. *Cognition and Instruction, 18*(3), 349-422.

Solomon, J., Duveen, J., & Scott, L. (1994). Pupils' images of scientific epistemology. *International Journal of Science Education, 16*(3), 361-373.

Solomon, J., Scott, L., & Duveen, J. (1996). Large-scale exploration of pupils' understanding of the nature of science. *Science Education, 80*(5), 493-508.

Sousa, V. D., Driessnack, M., & Mendes, I. A. C. (2007). An overview of research designs relevant to nursing: Part 1: quantitative research designs. *Revista latino-americana de enfermagem*, *15*(3), 502-507.

Southerland, S.A., Settlage, J., & Brickhouse, N. (2014). *Supporting the implementation of the Next Generation Science Standards (NGSS) through research: Accountability*. National Association for Research in Science Teaching opinion paper. Retrieved from

https://narst.org/ngsspapers/accountability.cfm.

Spencer, H. (1884). *What knowledge is of most worth*. Volume 3, Issue 138 of Elzevir library, University of Michigan.

Tabachnick, G., & Fidell, L. (2001). *Using Multivariate Statistics: 4th Edition*. Boston, MA: Allyn and Bacon.

Thomas, G., & Durant, J. (1987). Why should we promote the public understanding of science? In M. Shortland (Ed.), *Science literacy papers* (pp.1-14). Oxford, UK: Department for External Studies, University of Oxford.

Virginia Department of Education (2017). *Enrollment & Demographics*. Retrieved from: www.doe.virgnia.gov/statistics_reports/enrollment/index.shtml.

Wandersee, J.H., Mintzes, J.J., & Novak, J.D. (1994). Research on alternative conceptions in science. In D. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 177-210). New York: Macmillan.

Wenning, C. (2006). Assessing nature-of-science literacy as one component of scientific literacy. *Journal of Physics Teacher Education Online, 3*(4), 3-14.

West, S.G., Finch., J.F., Curran, P,J. (1995). Structural equation models with nonnormal variables: problems and remedies. In Hoyle, R.H. (eds.) Structural equation modeling: Concepts, issues and applications. Newbery Park, CA: Sage.

Wideen, M. F., O'Shea, T., Pye, I., & Ivany, G. (1997). High-stakes testing and the teaching of science. *Canadian Journal of Education*, 22(4), 428 – 444.

Wynne, B. (1991). Knowledges in context. *Science, Technology, & Human Values*, 16(1), 111–121.

Zepeda, C. D., Richey, J. E., Ronevich, P., & Nokes-Malach, T. J. (2015). Direct instruction of

metacognition benefits adolescent science learning, transfer, and motivation: An in vivo study. *Journal of Educational Psychology*, *107*(4), 954.

Ziman, J. (1991). Public understanding of science. *Science, Technology, & Human Values*, *16*(1), 99-105.

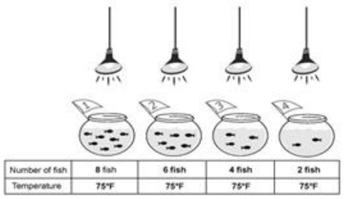
Zusho, A., Pintrich, P. R., & Coppola, B. (2003). Skill and will: The role of motivation and cognition in the learning of college chemistry. *International Journal of Science Education*, *25*(9), 1081-1094.

SLA-D1

SCIENCE SURVEY

***Directions*:** For each question, select one (1) answer and fill in corresponding circle on the separate bubble sheet.

1. A student is interested in the behavior of fish. He has 4 fish bowls and 20 goldfish. He puts 8 fish in the first bowl, 6 fish in the second bowl, 4 fish in the third bowl, and 2 fish in the fourth bowl. He places each fish bowl under light, he keeps the temperature at 75^{o} for all four bowls, and he observes the behavior of the fish.



What can the student find out from doing this experiment?

A. If the number of fish in the fish bowl affects the behavior of fish
B. If the temperature of the fish bowl affects the behavior of fish
C. If the temperature of the fish bowl and the amount of light affect the behavior of fish
D. If the number of fish, the temperature, and the amount of light affect the behavior of fish

2. A student finds a website created by the "No Homework Committee." He wants to find out the reasons for and against assigning homework to students**.** Is this a trustworthy source of information?

A. Yes. This group is against homework and knows all of the arguments.
B. Yes. Information on websites is always balanced and correct.
C. No. This group might give more attention to arguments against homework.
D. No. This group is probably not very good at arguing for or against homework.

3. In a town, 40% of the people have a certain illness. Fifty percent of the population is female. What percent of those who are ill are female?

A. 10% B. 20% C. 50% D. Cannot tell from the information given

4. Read the four questions below. Which question would be best answered by using scientific methods or instruments?

A. How much ice cream is sold each year?
B. How did the first people make ice cream?
C. How was ice cream kept cold before freezers?
D. How many calories are in a scoop of ice cream?

5. A group of students is making paper airplanes. They think that the kind of paper and the design of the airplane may affect how far each paper airplane flies. The students first test if the kind of paper affects how far the airplane flies. They make several airplanes out of different kinds of paper, using the same design. Why is it important that all the airplanes have the same design?

A. By using the same design, the students can learn about both the effect of the design and the effect of the paper.
B. By using the same design, the students can learn about the effect of the design.
C. If they do not use the same design, the students cannot learn about the effect of the paper.
D. It is **not** important for the airplanes to have the same design because the students are not testing the effect of the design.

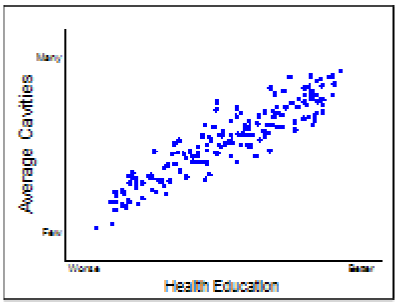
6. Andrew, Lisa, and Min each worked on ten new math problems, every day for a week. The table below shows the number of correct answers for each student.

	Mon	Tue	Wed	Thu	Fri
Andrew	4	5	8	6	5
Lisa	3	8	6	5	8
Min	4	5	5	6	9

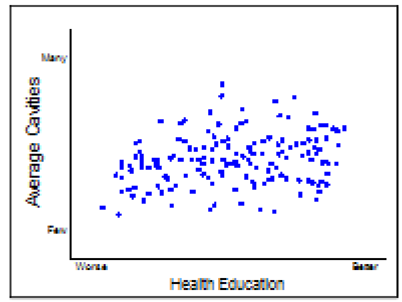
How many days did Lisa have the highest number correct?
A. 1
B. 2
C. 3
D. 4

7. Look at the graphs below. Which one shows that the average number of cavities per person is lower in countries with better health education?

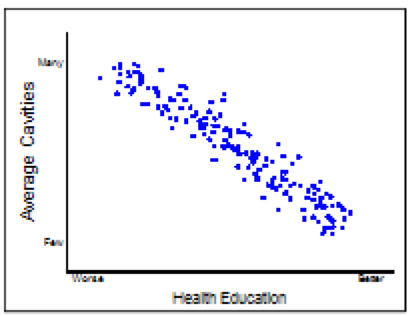
A. Graph A



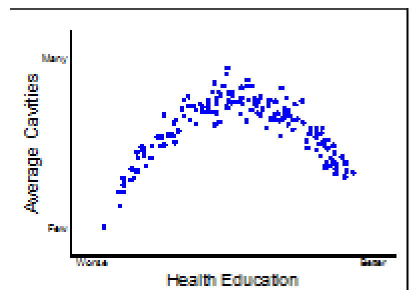
C. Graph C



B. Graph B



D. Graph D



8. You receive an email that claims "People who sleep with books under their pillows get better grades through osmosis."
Osmosis is a scientific process. Some molecules pass through a semi-permeable layer while larger molecules are blocked. It is used to get pure water from salt water. It is used for other processes too.

Based on this claim, should you sleep with your books to get better grades?

A. Yes, if you use a semi-permeable pillow.
B. Yes, because osmosis is a known scientific process.
C. No, because books under the pillow disrupt sleep.
D. No, because osmosis is not related to grades.

9. A state records the number of car crashes each year. In one year, there were 1,056 car crashes on 4- or 6-lane highways. In the same year there were 589 car crashes on 2-lane highways. The governor concludes that driving on 4- or 6-lane highways is more dangerous than driving on 2-lane highways. What do you think about this conclusion?

A. Correct. Highways that have 4 or 6 lanes are known to be more dangerous.
B. Incorrect. The number of accidents on 2 lane highways might be increasing.
C. Correct. States have more highways with 4 or 6 lanes than with 2 lanes.
D. Incorrect. He did not consider numbers of cars using each kind of highway.

10. On February 2 in the U. S., many news reporters travel to local zoos to watch the behavior of a groundhog. Some people say that if the groundhog leaves his burrow and stays outside, spring will come sooner; if he runs back inside, there will be six more weeks of winter. This year the groundhog ran back inside.

What is the correct scientific observation about the occurrence?

A. The groundhog returned to his burrow.

B. There will be six more weeks of winter.

C. The groundhog was frightened by the reporters

D. Groundhogs have inborn weather reading abilities.

11. Arturo's parents want him to get better grades in school. His mother read a research study on the topic. After reading the study, she decided that from now on Arturo needed to be in bed by 9 PM. Which of these studies did Arturo's mother read?

A. When students go to bed by 9 pm they are less tired in school.

B. Students who get good grades are more alert when at school.

C. When students go to bed by 9 pm their school grades improve.

D. Students who go to bed earlier have more energy the next day.

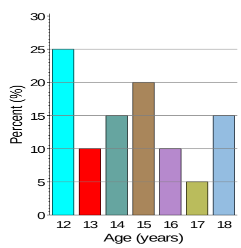
12. What percent of the sample of people shown in the graph is older than 15 years?

A. 20%

B. 30%

C. 40%

D. 50%



13. Sue wants to find out what growing conditions might affect the length of bean seedlings. She places a bean wrapped in moist tissue paper in each of ten identical test tubes. She puts five tubes in a rack in a sunny window. She puts the other five tubes in a dark refrigerator. Sue measures the lengths of the bean seedlings in each group of test tubes after one week.

Look at the variables listed below. Which variables was Sue testing to see how they affect the length of the bean seedlings?

A. Temperature and moisture

B. Moisture and length of test tube

C. Light and temperature

D. Light and amount of time

14. A country has a high number of decayed teeth (cavities) per person. Which question about tooth decay can only be answered with scientific experiments?

A. Do the men in this country have more tooth decay than women?

B. Would putting vitamin D in the water supply affect tooth decay?

C. Has the number of decayed teeth increased in the past 10 years?

D. Is tooth decay more common in some parts of the country than others?

15. The principal of Riley Middle School wants to remove candy and soda (pop) vending machines. In their place, she wants to put in healthy food machines. She wants to know what her students will think of these changes. What would be the best way to get an accurate answer to this question?

    A. Give a survey to all students who play on sports teams.

    B. Give a survey to all students who attend a health fair.

    C. Give a survey to every 20th student on a list of all students.

    D. Give a survey to all students who use the vending machines.

16. Billy and George raced their bikes five times. Billy made a graph of their a

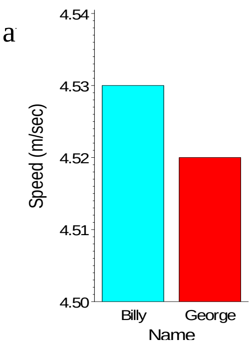
    What can you reasonably conclude from the graph?

    A. George is a much worse bike rider than Billy.

    B. Billy is a lot faster because his average speed is much higher than George's speed.

    C. They are about the same because there is not much difference in their speed.

    D. Billy won all five races.



17. A town council wants to protect cats from being hit by cars on town streets. They decide to adopt a policy of banning all pet cats from being outside. Which of the following is likely to be an <u>unintended</u> effect of this policy?

    A. A rise in mouse populations

    B. A drop in pet cat populations

    C. An increase in car accidents

    D. An decrease in dog ownership

18. A TV weather report said, "New Jersey is heading for a severe water shortage!" Which type of evidence listed below would be the most important to support this claim?

    A. Average weekly rainfall

    B. Average weekly temperature

    C. Water levels at high tide

    D. Number of cloudy days

19. A science journal publishes a study about effects of diet in rats. For six weeks, the scientists fed only dog food to 60 male rats. They fed normal rat food to another 60 male rats for the same six weeks. At the end of the six weeks, they counted the number of rats in each group that had developed dark spots. The results are in the table below.

	Rats That Developed Dark Spots	Rats That Did Not Develop Dark Spots	TOTAL
Fed dog food	42	18	60
Fed normal rat food	16	44	60

Which conclusion is suggested by the results in this table?

A. It is good to feed dog food to rats.

B. Rats with dark spots prefer dog food.

C. Type of food is linked to dark spots in rats.

D. Diet has nothing to do with rats getting dark spots.

Appendix B

SLA-MB

Value of Science

Directions: For each item below select the number that best describes how much you agree or disagree with each statement. These are your opinions and there are no "right" or "wrong" answers.

1.	In general, I find working on science assignments	① Very boring	②	③	④	⑤ Very Interesting (Fun)
2.	Compared to most of your other activities, how useful is what you learn in science?	① Not at all useful	②	③	④	⑤ Very useful
3.	For me, being good in science is	① Not at all important	②	③	④	⑤ Very importan t
4.	Compared to most of your other activities, how important is it for you to be good at science?	① Not at all important	②	③	④	⑤ Very importan t
5.	How much do you like doing science?	① Not at all	②	③	④	⑤ Very much
6.	Some things that you learn in school help you do things better outside of class, that is, they are useful. For example, learning about plants might help you grow a garden. In general, how useful is what you learn in science?	① Not at all useful	②	③	④	⑤ Very useful

What I Can Do in Science

Directions. Select the number that best describes how much you agree with each statement below. These are your opinions and there are no "right" or "wrong" answers.

		Strongly Disagree	Disagre e	Somewhat	Agree	Strongly Agree
7.	I know when to use science to answer questions.	①	②	③	④	⑤
8.	I can use science to make decisions about my daily life.	①	②	③	④	⑤
9.	I know how to use the scientific method to solve problems.	①	②	③	④	⑤
10.	It is easy for me to tell the difference between scientific findings and advertisements.	①	②	③	④	⑤
11.	When I do my work in science class, I am able to find the important ideas.	①	②	③	④	⑤

12.	I can use math to answer scientific questions.	①	②	③	④	⑤
13.	I can tell the difference between observations and conclusions in a story.	①	②	③	④	⑤
14.	It is easy for me to make a graph of my data.	①	②	③	④	⑤

What I Believe about Science

Directions. Select the number that best describes how much you agree with each statement below. These are your opinions and there are no "right" or "wrong" answers. Use the attached bubble sheet to record your answers.

		Strongly Disagree				Strongly Agree
15.	Everybody has to believe what scientists say.	①	②	③	④	⑤
16.	All questions in science have one right answer.	①	②	③	④	⑤
17.	Scientific knowledge is always true	①	②	③	④	⑤
18.	In science, you have to believe what the science books say about stuff.	①	②	③	④	⑤
19.	The most important part of doing science is coming up with the right answer.	①	②	③	④	⑤
20.	Whatever the teacher says in science class is true.	①	②	③	④	⑤
21.	Scientists pretty much know everything about science; there is not much more to know.	①	②	③	④	⑤
22.	If you read something in a science book, you can be sure it's true.	①	②	③	④	⑤
23.	Once scientists have a result from an experiment that is the only answer.	①	②	③	④	⑤
24.	Scientists always agree about what is true in science	①	②	③	④	⑤
25.	Only scientists know for sure what is true in science.	①	②	③	④	⑤

Appendix C

Modified Attitudes Toward Science Inventory (mATSI)

Directions: Please circle the number that most closely represents your level of agreement to each of the following statements.

	Strongly Disagree	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**
1. Science teachers make science interesting.	1	2	3	4	5
2. Science teachers present material in a clear and understandable way.	1	2	3	4	5
3. Science teachers are willing to give us individual help.	1	2	3	4	5
4. When I hear the word "science," I have a feeling of dislike.	1	2	3	4	5
5. I feel tense when someone talks to me about science.	1	2	3	4	5
6. It makes me nervous to even think about doing science.	1	2	3	4	5
7. It scares me to have to take science class.	1	2	3	4	5
8. I have a good feeling toward science.	1	2	3	4	5
9. Science is useful for solving the problems of everyday life.	1	2	3	4	5
10. Most people should study some science.	1	2	3	4	5
11. Science is helpful in understanding today's world.	1	2	3	4	5
12. Science is of great importance to a country's development.	1	2	3	4	5
13. It is important to know science in order to get a good job.	1	2	3	4	5
14. I do not do very well in science.	1	2	3	4	5
15. Science is easy for me.	1	2	3	4	5
16. I usually understand what we are talking about in science.	1	2	3	4	5
17. No matter how hard I try, I cannot understand science.	1	2	3	4	5

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
18. I often think, “I cannot do this,” when a science assignment seems hard.	1	2	3	4	5
19. Science is something which I enjoy very much.	1	2	3	4	5
20. I would like to do some reading in science which has not been assigned to me.	1	2	3	4	5
21. Sometimes I read ahead in our science book.	1	2	3	4	5
22. I like the challenge of science assignments.	1	2	3	4	5
23. It is important to me to understand the work I do in the science class.	1	2	3	4	5
24. Science is one of my favorite subjects.	1	2	3	4	5
25. I have a real desire to learn science.	1	2	3	4	5

Appendix D

STEM Career Interest Survey (STEM-CIS)

Directions: Please circle the number that most closely represents your level of agreement to each of the following statements.

		Strongly Disagree				**Strongly Agree**
1.	I am able to get a good grade in my science class.	1	2	3	4	5
2.	I am able to complete my science homework.	1	2	3	4	5
3.	I plan to use science in my future career.	1	2	3	4	5
4.	I will work hard in my science classes.	1	2	3	4	5
5.	If I do well in science classes, it will help me in my future career.	1	2	3	4	5
6.	My parents would like it if I choose a science career.	1	2	3	4	5
7.	I am interested in careers that use science.	1	2	3	4	5
8.	I like my science class.	1	2	3	4	5
9.	I have a role model in a science career.	1	2	3	4	5
10.	I would feel comfortable talking to people who work in science careers.	1	2	3	4	5
11.	I know of someone in my family who uses science in their career.	1	2	3	4	5

Middle School General Science Knowledge Survey

Directions

Read each question and choose the best answer. Then circle the letter that corresponds with the answer you have chosen.

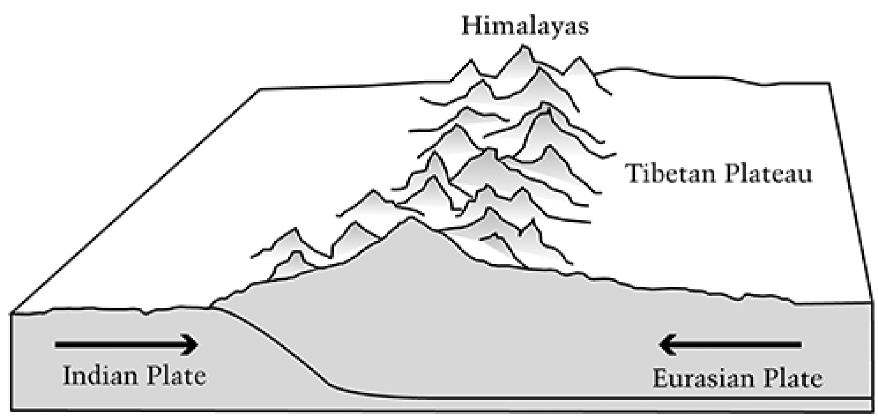
1. Which characteristic is shared by all cells?

A. They need energy.
B. They reproduce sexually.
C. They make their own food.
D. They move from place to place.

2. Kelly slides a flat rock across the smooth ice of a frozen pond. The rock slows down after several seconds. What causes the rock to slow down?

A. The thickness of the ice
B. The temperature of the air above the ice
C. The force of friction between the ice and the rock
D. The gravitational force between the ice and the rock

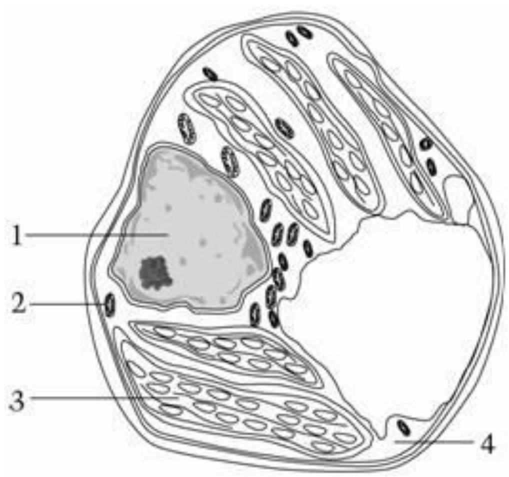
The diagram below shows the collision of two tectonic plates in Asia.



3. What is a result of this collision?

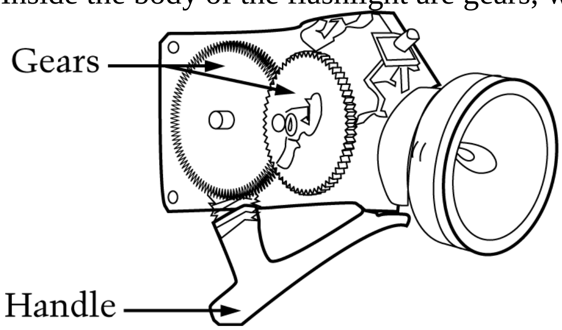
A. Volcanoes erupt periodically.
B. The Tibetan Plateau slowly sinks.
C. The Himalayas increase in height each year.
D. Glaciers on the Tibetan Plateau melt.

4. In the picture of a cell below, which label indicates the part of the cell that contains most of the cell's genetic material?



A. 1
B. 2
C. 3
D. 4

The flashlight shown below has no batteries. It is operated by squeezing and letting go of the handle. Inside the body of the flashlight are gears, which are shown below

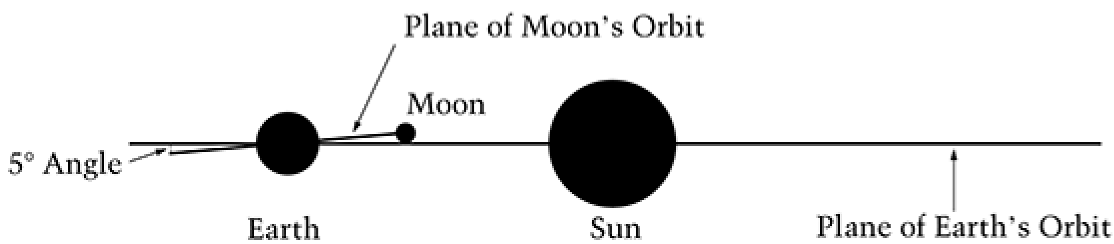


5. Which sequence best identifies the energy transfers that take place within the flashlight to produce light?

A. Kinetic ⟶ electrical ⟶ light
B. Kinetic ⟶ chemical ⟶ light
C. Chemical ⟶ kinetic ⟶ light
D. Chemical ⟶ electrical ⟶ light

Earth revolves around the Sun, and the Moon revolves around Earth. The Moon's orbital path is sometimes above and sometimes below the plane of Earth's orbit, as shown in the diagram below.

VIEW OF SUN, EARTH, AND MOON ALONG PLANE OF EARTH'S ORBIT



6. What would happen if Earth's orbit and the Moon's orbit were in the same plane?

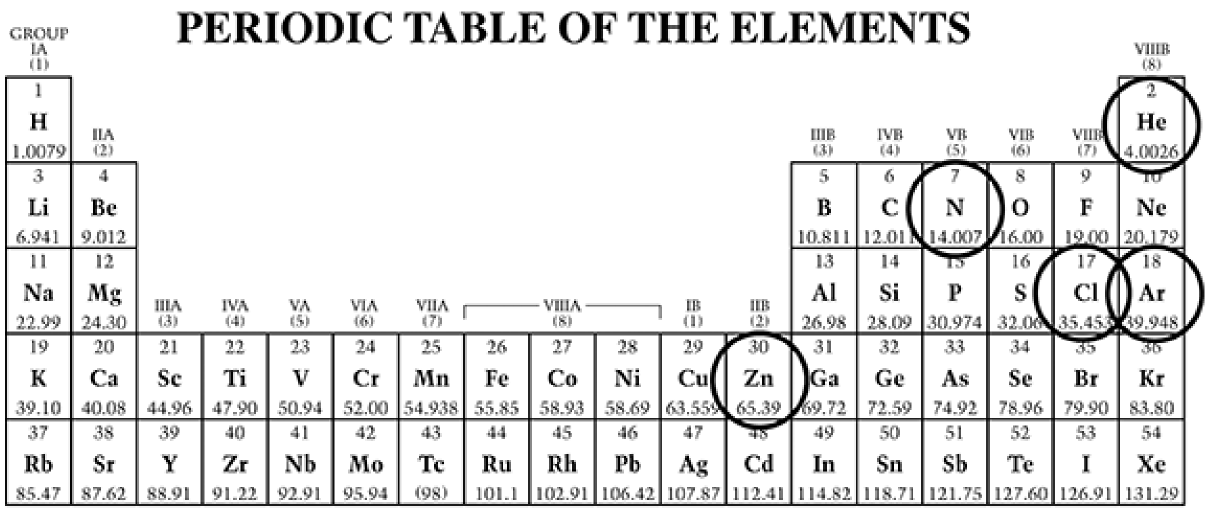
A. Eclipses would occur every month.
B. The Moon would not have phases.
C. All sides of the Moon would be visible from Earth.
D. The same side of the Moon would always face the Sun.

7. A person produces two sound waves with a flute, one immediately after the other. Both sound waves have the same pitch, but the second one is louder. Which of the following properties is greater for the second sound wave?

A. Frequency
B. Amplitude
C. Wavelength
D. Speed in air

8. Which of the following is most consistent with the modern theory of evolution?

A. Parents pass their physical traits to their offspring; those offspring with traits that help them survive in the environment are able to reproduce.
B. Parents change their physical traits in order to survive in the environment, then those parental traits are passed to their offspring.
C. Life on this planet came from another planet far out in space.
D. Living organisms have not changed for hundreds of millions of years.



PERIODIC TABLE OF THE ELEMENTS

GROUP IA (1)	IIA (2)	IIIA (3)	IVA (4)	VA (5)	VIA (6)	VIIA (7)	VIIIA (8)			IB (1)	IIB (2)	IIIB (3)	IVB (4)	VB (5)	VIB (6)	VIIB (7)	VIIIB (8)
1 H 1.0079																	2 He 4.0026
3 Li 6.941	4 Be 9.012											5 B 10.811	6 C 12.011	7 N 14.007	8 O 16.00	9 F 19.00	10 Ne 20.179
11 Na 22.99	12 Mg 24.30											13 Al 26.98	14 Si 28.09	15 P 30.974	16 S 32.06	17 Cl 35.453	18 Ar 39.948
19 K 39.10	20 Ca 40.08	21 Sc 44.96	22 Ti 47.90	23 V 50.94	24 Cr 52.00	25 Mn 54.938	26 Fe 55.85	27 Co 58.93	28 Ni 58.69	29 Cu 63.559	30 Zn 65.39	31 Ga 69.72	32 Ge 72.59	33 As 74.92	34 Se 78.96	35 Br 79.90	36 Kr 83.80
37 Rb 85.47	38 Sr 87.62	39 Y 88.91	40 Zr 91.22	41 Nb 92.91	42 Mo 95.94	43 Tc (98)	44 Ru 101.1	45 Rh 102.91	46 Pb 106.42	47 Ag 107.87	48 Cd 112.41	49 In 114.82	50 Sn 118.71	51 Sb 121.75	52 Te 127.60	53 I 126.91	54 Xe 131.29

Element	Symbol
Argon	Ar
Chlorine	Cl
Helium	He
Nitrogen	N
Zinc	Zn

9. Based on its location on the partial periodic table shown above, which element would you predict has chemical properties that are most similar to argon (Ar)?

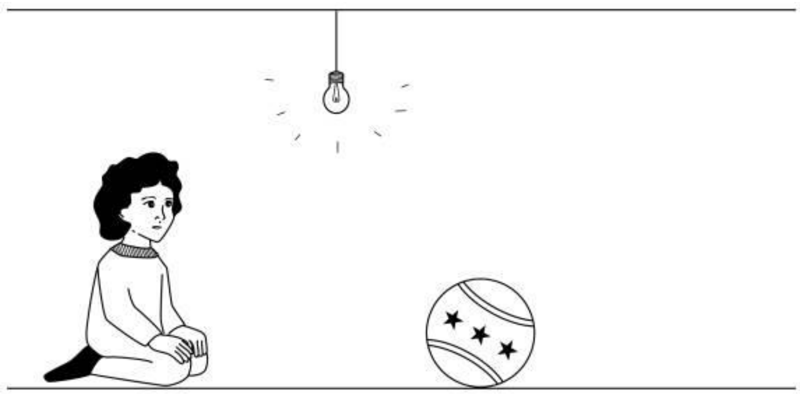
A. Chlorine (Cl)
B. Helium (He)
C. Nitrogen (N)
D. Zinc (Zn)

10. It is important to protect air quality because —

A. Storms worsen as air pollution decreases
B. Acid rain is caused by air pollution
C. Wind currents change when the air is polluted
D. Energy produced by the Sun decreases when air is polluted

11. All of these can be inherited by people EXCEPT —

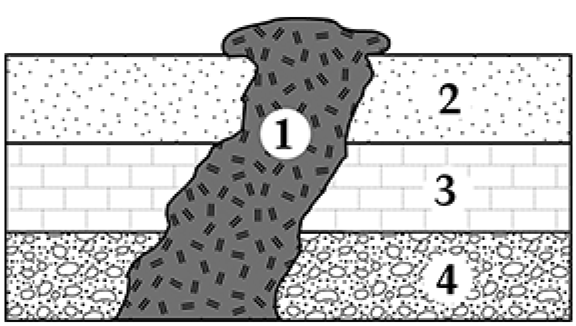
A. Height
B. Eye color
C. Blood type
D. Language



12. In the figure above, which of the following is the pathway of light that allows the child to see the ball?

A. Light bulb → child's eyes → ball
B. Child's eyes → light bulb → ball
C. Ball → light bulb → child's eyes
D. Light bulb → ball → child's eyes

The diagram below shows a cross section of rock formations.



13. Which rock formation was formed most recently?

A. 1
B. 2
C. 3
D. 4

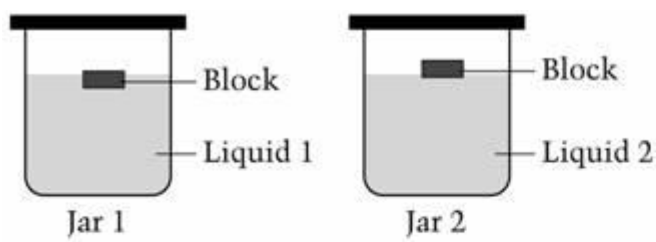
14. Which energy transformation occurs first in a coal-burning power plant?

A. Chemical energy to thermal energy
B. Thermal energy to mechanical energy
C. Thermal energy to electrical energy
D. Mechanical energy to electrical energy

15. Which characteristic is used to classify frogs into a different phylum from squid, snails, and jellyfish?

A. Frogs are predators.
B. Frogs breathe oxygen.
C. Frogs have backbones.
D. Frogs live on land.

16. Look at the two pictures below. They show what happened when two solid blocks were " each put in a jar containing a liquid. Based just on what you can see in the pictures, what can you say about the blocks and the jars?



A. The liquid in the jars must be water.
B. The block in jar 1 weighs more than the block in jar 2.
C. The block in jar 1 is floating lower in its liquid than is the block in jar 2.
D. The block in jar 1 must be made of metal and the block in jar 2 must be made of wood.

17. What property of water is most important for living organisms?

A. It is odorless.
B. It does not conduct electricity.
C. It is tasteless.
D. It is liquid at most temperatures on Earth.

18. Why does the leaf of a plant look green?

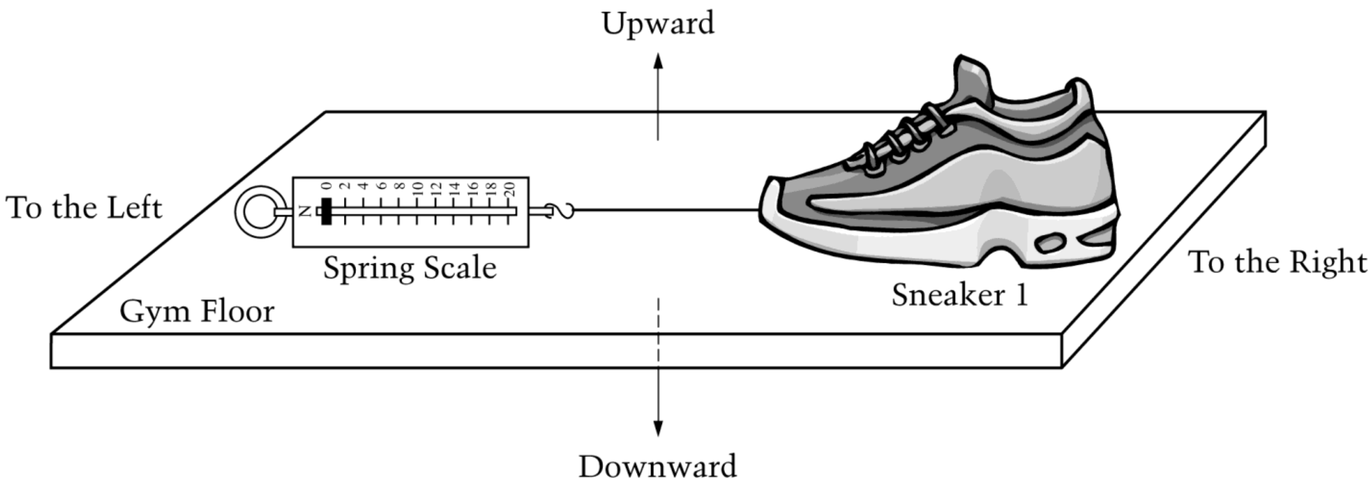
A. Because it absorbs green light
B. Because it reflects green light
C. Because it absorbs only yellow and blue light
D. Because it reflects a mixture of yellow and blue light

Meg designs an experiment to see which of three types of sneakers provides the most friction.

She uses the equipment listed below.

- Sneaker 1
- Sneaker 2
- Sneaker 3
- Spring scale

She uses the setup illustrated below and pulls the spring scale to the left

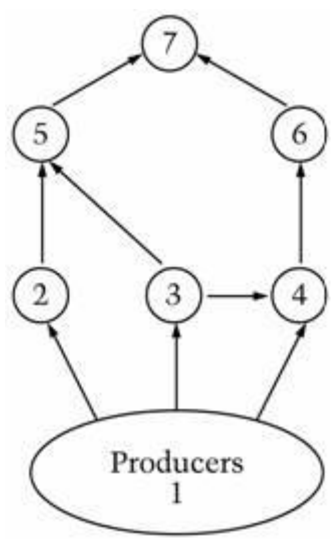


19. In what direction does the force of friction act?

A. To the left
B. To the right
C. Upward
D. Downward

20. **If you breathe on a mirror, part of the mirror clouds up. What are you actually seeing when you see the mirror cloud up?**

A. Water droplets that formed from cooled water vapor in your breath
B. Carbon dioxide that you are breathing out from your lungs
C. Oxygen that you are breathing out from your lungs
D. Cooled nitrogen in the air around you



21. **The diagram above shows a food web in a large park. Each circle represents a different species in the food web. Which of the organisms in the food web could be referred to as primary consumers?**

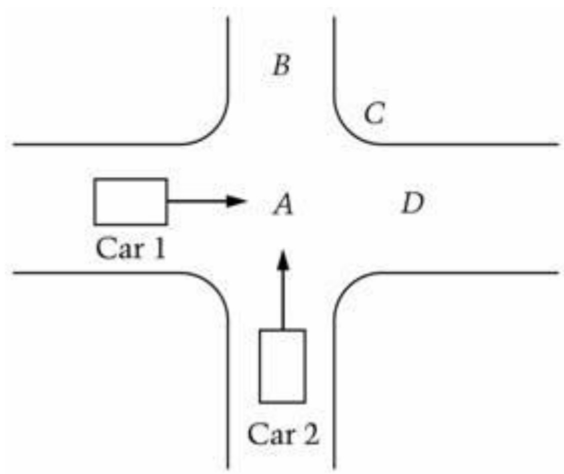
A. 7 only
B. 5 and 6 only
C. 2, 3, and 4 only
D. 2, 5, and 7 only

22. **Which of the following is an example of genetic engineering?**

A. Growing a whole plant from a single cell.
B. Finding the sequences of bases in plant DNA.
C. Inserting a gene into plants that makes them resistant to insects.
D. Attaching the root of one type of plant to the stem of another type of plant.

23. If air pollution causes the rain that falls on this pond to become much more acidic, after two years how will this acidity affect the living things in this pond?

A. There will be more plants and animals because the acid is a source of food.
B. There will be fewer plants and animals because the acid will dissolve many of them.
C. There will be fewer plants and animals because many of them cannot survive in water with high acidity.
D. There will be more plants and animals because the acid will kill most of the disease-causing microorganisms.



24. Two identical cars travel at 45 miles per hour toward the center of the intersection (point A, as shown above) with equal force. The cars collide at the intersection. If after they collide the cars stick to each other and move together, they will come to rest closest to

A. point A
B. point B
C. point C
D. point D

25. During which phase does the Moon receive sunlight only on the side facing away from Earth?

A. Full Moon
B. New Moon
C. Waning gibbous
D. Waxing gibbous

Appendix F

High School General Science Knowledge Survey

Directions

Read each question and choose the best answer. Then circle the letter that corresponds with the answer you have chosen.

1. **What is the correct order for the levels of organization in living systems from the simplest to the most complex?** (Note that not all levels of organization are included.)

   A. Elements ⟶ molecules ⟶ cells ⟶ tissues ⟶ organs

   B. Molecules ⟶ tissues ⟶ cells ⟶ organs ⟶ organisms

   C. Molecules ⟶ elements ⟶ tissues ⟶ organs ⟶ organisms

   D. Cells ⟶ organs ⟶ tissues ⟶ organisms ⟶ molecules

2. **Coal, petroleum, and natural gas found underground in certain parts of Earth are primarily formed from which process?**

   A. Decay of radioactive elements
   B. Collision of tectonic plates in earthquakes
   C. Transformation of dead plants and animals under heat and pressure
   D. Intrusion of water into the soil that breaks up rocks and minerals

3. **An unusual type of fossil clam is found in rock layers high in the Swiss Alps. The same type of fossil clam is also found in the Rocky Mountains of North America. From this, scientists conclude that ---**

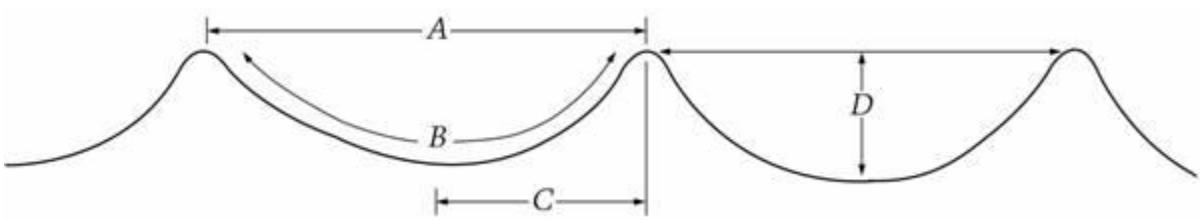
   A. glaciers carried the fossils up the mountains
   B. the Rocky Mountains and the Swiss Alps are both volcanic in origin
   C. clams once lived in mountains, but have since evolved into sea-dwelling creatures
   D. the layers of rocks in which the fossils were found are from the same geologic age

4. **What two gases make up most of the Earth's atmosphere?**

   A. Hydrogen and oxygen
   B. Hydrogen and nitrogen
   C. Oxygen and carbon dioxide
   D. Oxygen and nitrogen

5. A person produces two sound waves with a flute, one immediately after the other. Both sound waves have the same pitch, but the second one is louder. Which of the following properties is greater for the second sound wave?

A. Frequency
B. Amplitude
C. Wavelength
D. Speed in air



6. The figure above shows some ocean waves. Which of the labeled distances represents the wavelength?

A. *A*
B. *B*
C. *C*
D. *D*

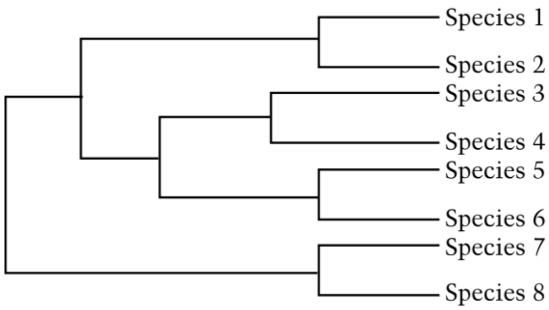
7. Which particle is a negatively charged ion?

A. Hydrogen (H) with 1 proton, 0 neutrons, and 1 electron
B. Sodium (Na) with 11 protons, 12 neutrons, and 10 electrons
C. Chlorine (Cl) with 17 protons, 18 neutrons, and 18 electrons
D. Magnesium (Mg) with 12 protons, 12 neutrons, and 12 electrons

8. A geneticist studying fruit flies hypothesizes that short wings are a recessive trait coded for by a single gene. Which observation is most likely to have led her to form this hypothesis?

A. Flies have wing lengths ranging from very long to very short.
B. Flies with long wings are less likely to survive.
C. Flies with long wings can produce offspring with short wings.
D. Flies with short wings prefer to mate with flies with long wings.

Scientists are studying the evolutionary history of a group of plants in the United States, and they develop an evolutionary tree, as shown below

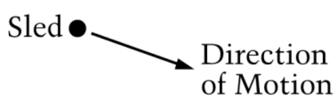


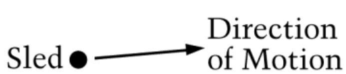
9. Which statement can be inferred from the evolutionary tree?

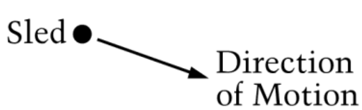
A. Species 1 is most closely related to Species 8.
B. Species 2 is most closely related to Species 3.
C. Species 3 is most closely related to Species 7.
D. Species 5 is most closely related to Species 6.

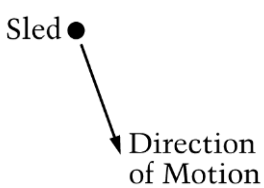
Two dogs pull on a flat-bottom sled with forces of equal magnitude in the directions indicated by the arrows below. The dot represents the sled.

10. Which arrow best represents the direction of motion of the sled?

A. 

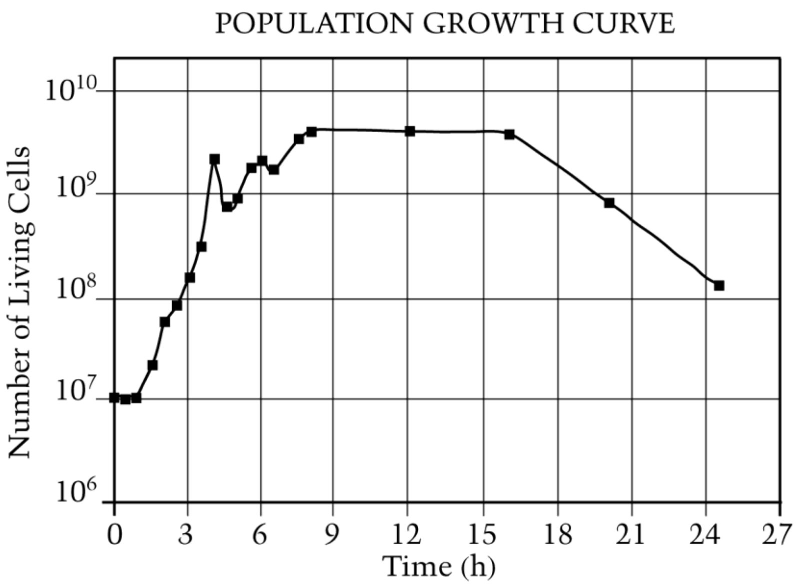
B. 

C. 

D. 

A scientist studied the growth rate of a species of bacterium. The scientist introduced some of the bacteria into a flask of nutrient-rich solution and monitored the growth of the bacterial population by measuring the number of living cells in the solution.

The graph below shows the growth of the bacterial population over time in hours (h).

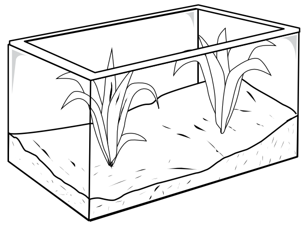


11. Over which time period did the number of living bacteria increase at the greatest rate?

A. Between hours 0 and 1
B. Between hours 1 and 8
C. Between hours 8 and 16
D. Between hours 16 and 24

12. What is the main reason that water has the ability to dissolve many different substances?

A. Water has a lower molecular mass than many substances.
B. Water molecules attract ions and the charged parts of molecules.
C. Water molecules are larger than the ions or molecules they dissolve.
D. Water is more dense in the liquid phase than in the solid phase.



13. A student constructs several terrariums like the one shown. Each terrarium is exposed to a different amount of sunlight each day. In order to determine the ideal amount of sunlight, which of the following variables must be held constant?

A. Type of plants
B. Growth rate of plants
C. Wavelengths of sunlight
D. Amount of sunlight received

	Charcoal	Carbon Dioxide
Formula	C	CO_2
State at Room Temperature	Solid	Gas
Soluble in Water	No	Yes
Combustible in Air	Yes	No

14. Based on the information in the table above, which is a reasonable hypothesis regarding elements and their compounds?

A. An element retains its physical and chemical properties when it is combined into a compound.
B. When an element reacts to form a compound, its chemical properties are changed but its physical properties are not.
C. When an element reacts to form a compound, its physical properties are changed but its chemical properties are not.
D. Both the chemical and physical properties of a compound are different from the properties of the elements of which it is composed.

15. Which of the following observations about a certain pure solid would indicate most strongly that the solid is ionic?

A. Its water solution is a good conductor of electricity.
B. It is composed of small white crystals.
C. It has a density greater than 1.0 gram/cm ^{3}.
D. It has a high melting point.

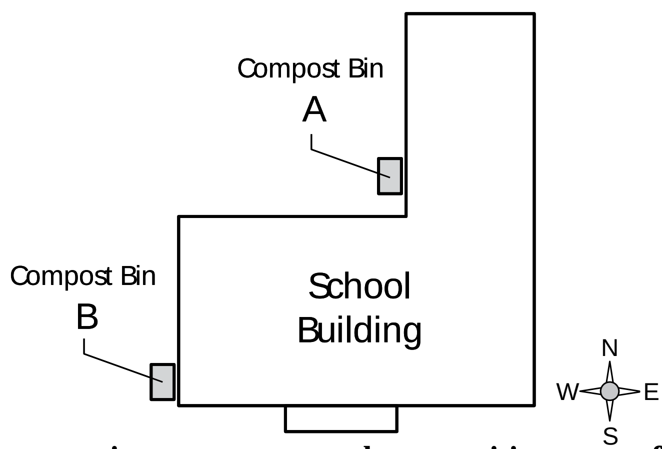
A student took a sample of water from a pond and examined it under a microscope. She identified several species of protozoans, including two species of *Paramecium* that are known to eat the same food. The student decided to examine the water sample every day for a week. She added food for the *Paramecia* each day and counted the number of each species. Her findings are summarized in the table below.

NUMBER OF *PARAMECIA* IN POND WATER SAMPLE

Day	Species *S*	Species *T*
1	50	50
2	60	80
3	100	90
4	150	60
5	160	50
6	160	30
7	160	20

16. Which of the following can be correctly concluded from the data?

A. Species *S* is the food for species *T*.
B. Species *T* is more common than species *S*.
C. Species *S* is a more successful competitor than species *T*.
D. Species *T* is a more successful competitor than species *S*.



17. As part of an experiment to measure decomposition rates of different materials, students put food scraps from the cafeteria in compost bin A and leaves and grass clippings in compost bin B for six weeks. Students in first period measured the temperature in bin A, and students in sixth period measured the temperature in bin B. What is the greatest error in the students' experimental design?

A. There are too many uncontrolled variables in the experiment.
B. Temperature is the only dependent variable in the experiment.
C. The materials chosen decompose too rapidly.
D. The students put equal masses of materials in each bin.