

MEASURING ELEMENTARY EDUCATORS' UNDERSTANDING AND
READINESS FOR IMPLEMENTING A NEW FRAMEWORK
IN SCIENCE EDUCATION

by

Gustave Evan Nollmeyer

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

of

Doctor of Education

in

Education

MONTANA STATE UNIVERSITY
Bozeman, Montana

April, 2014

©COPYRIGHT

by

Gustave Evan Nollmeyer

2014

All Rights Reserved

TABLE OF CONTENTS

1. INTRODUCTION

An Overview of Science Education	1
Defining the Study of Science	1
Nationwide Reforms in Science Education	5
Current Efforts to Revise National Science Education Standards	7
Statement of the Problem.....	8
Research Questions	10
Limitations and Delimitations.....	10
Definitions of Terms	11
Significance of the Study	13

2. LITERATURE REVIEW

Introduction.....	14
The Current State of Elementary Science.....	14
The History of Science Education	16
Origins of Science Education (Before Circa 1880)	17
Early Structuring (Circa 1880 to 1920)	19
Restructuring Efforts (Circa 1920 to 1950)	22
National Reforms and Elementary Science (Circa 1950 to 1980).....	25
National Science Standards (1980 to 2000).....	29
Science Education in the 21 st Century (2000 to 2014)	34
The Science and Psychology of Learning Science	38
How Children Know and Understand.....	38
How Children Generate and Evaluate.....	42
How Children Understand the Construction of Scientific Knowledge.....	43
How Children Participate in Science	44
Revisions to Teacher Education.....	46
<i>Next Generation Science Standards</i> ' Connection to The Framework	50
Scientific & Engineering Practices	52
Crosscutting Concepts	53
Disciplinary Core Ideas.....	55
Integration and The Framework.....	56
Relevant Aspects of Integration.....	57
The Intradisciplinary and Interdisciplinary Nature of Science	63
The Framework Features that Promote Integration	64
Instrument Design.....	67
Describing what is Being Measured	67
Ensuring Validity	68
Ensuring Reliability	71

TABLE OF CONTENTS – CONTINUED

Themes for The Framework.....	73
Scientific & Engineering Practices	74
Crosscutting Concepts	75
Disciplinary Core Ideas.....	76
Integration of Three Dimensions	76
Best Practices in Science Education	77
Connections to Common Core.....	78
The Constructs of Understanding and Readiness	79
Chapter Summary	80
3. RESEARCH METHODOLOGY	
Introduction.....	82
Participants.....	83
Instrument Development.....	84
Creating the Item Pool	87
Expert Review.....	83
Piloting the Draft Instrument	92
Administering the <i>NFSE-STUR</i>	94
Data Analysis Methods.....	95
4. RESULTS	
Introduction.....	97
Expert Review.....	97
Pilot Study.....	113
Validation Study	117
Introduction.....	117
Participant Demographics.....	118
Descriptive Analysis	120
<i>Understanding</i> Scale.....	125
Exploratory Analysis	125
Whole Scale and Factor Reliability Analysis	131
Confirmatory Factor Analysis.....	132
<i>Readiness</i> Scale.....	134
Exploratory Analysis	134
Whole Scale and Factor Reliability Analysis	139
Confirmatory Factor Analysis.....	139
Post Hoc Analysis on Crosscutting Concepts Theme.....	141
Chapter Summary	142

TABLE OF CONTENTS - CONTINUED

5. DISCUSSION

Introduction.....143
Using the *NFSE-STUR* to Assess *Understanding* and *Readiness*.....144
Discussing the Latent Variables for the *Understanding* Scale (*NFSE-STU*).....145
Discussing the Latent Variables for the *Readiness* Scale (*NFSE-STR*).....151
Implications and Applications for Practice.....158
 The Uniqueness and Interdependence of the Three Dimensions.....158
 Connections to Common Core State Standards, Best Practice or Integration? ...159
 Implications for Professional Development.....160
Recommendations for Future Research161
Conclusions.....162

REFERENCES CITED.....165

APPENDICES182

 APPENDIX A: *Expert Review Draft: NFSE-STUR*183
 APPENDIX B: Expert Review Stage Two Power Point187
 APPENDIX C: *Pilot Draft: NFSE-STUR*.....190
 APPENDIX D: *Validation Draft: NFSE-STUR*194
 APPENDIX E: Expert Review Stage One Comments198
 APPENDIX F: *New Framework for Science Education:*
 Survey of Teacher Understanding206
 APPENDIX G: *New Framework for Science Education:*
 Survey of Teacher Readiness214

LIST OF TABLES

Table	Page
1. Descriptive Statistics for Expert Review of Initial Instrument Items.....	99
2. Item Changes Resulting from Expert Review	106
3. Demographic Characteristics of Pilot Study Participants, $N = 13$	113
4. Demographic Characteristics of Validation Study Participants, $N = 167$	118
5. Descriptive Statistics for <i>Understanding</i> and <i>Readiness</i> Scales	121
6. Rotated Factor Structure for the <i>NFSE-STUR Understanding</i> Scale	127
7. Whole Scale & Factor Reliability for <i>NFSE-STUR Understanding</i> Construct.....	132
8. Rotated Factor Structure for the <i>NFSE-STUR Readiness</i> Scale	136
9. Whole Scale and Factor Reliability for <i>NFSE-STUR Readiness</i> Construct	139
10. Comparing <i>Understanding</i> Scale Development with the Factor Model	147
11. Comparing <i>Readiness</i> Scale Development with the Factor Model	152

LIST OF FIGURES

Figure	Page
1. Unifying the Practices of Science & Engineering	53
2. Screenshot of Pilot Draft Item 8	86
3. Screenshot of Framework Quote for Pilot Draft Item 8	86
4. Screenshot of Item 23 on the Expert Review State One Survey	91
5. <i>Understanding</i> Scale Model Fit.....	133
6. <i>Readiness</i> Scale Model Fit	140
7. NGSS Logo Demonstrating the Intended Interdependence of the Three Dimensions	149

ABSTRACT

The NRC's (2012) report, *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*, ushered in a new era of science education. It is unclear how prepared elementary educators are for the framework. This study sought to establish measures for assessing inservice educators' self-reported *understanding* of the new framework and *readiness* to implement the ideas in their science instruction. Designing and validating an instrument to assess these constructs followed procedures established in the literature. First, literature on science education was examined to identify themes that could be used in constructing instrument items. This item pool was examined and modified through expert review. Next, the modified instrument was piloted with a small sample, $N = 13$, of inservice educators. After final adjustments to the instrument, it was used in a large scale validation study. Inservice elementary educators from four states, Montana, Idaho, Wyoming, and Utah participated in the validation study, $N = 167$. Since *understanding* and *readiness* were determined to assess separate constructs, the two were handled individually during statistical analyses. Exploratory analysis on both scales, *understanding* and *readiness*, revealed stable factor models that were further validated through confirmatory factor analysis. The internal consistency reliability of the scales were determined through Cronbach's Alpha. With solid statistical evidence, conclusions were drawn from the study. Each instrument could be used in similar contexts to measure elementary educators' *understanding* of or *readiness* to implement the new framework for science education. The unique factor structures of the two scales suggests important differences between *understanding* and *readiness*. These differences should inform professional development efforts.

CHAPTER ONE

INTRODUCTION

An Overview of Science EducationDefining the Study of Science

Any discussion about teaching and learning science must first endeavor to define science. This is a more challenging task than it first appears, yet a clear definition is vital because it guides a conceptualization of what it means to teach science and what it means to learn science. As a scientist Quinn (2009) explains that science is both a process of discovery and a collection of knowledge resulting from that process. This view is shared by the National Research Council (NRC) who suggest that it is the best way for all, not just scientists, to conceptualize science, “Science is both a body of knowledge that represents current understanding of natural systems and the process whereby that body of knowledge has been established and is being continually extended, refined, and revised” (National Research Council, 2007, p. 26). The NRC continues, pointing out that both elements are essential. The world of science cannot progress if one is lacking. However, all too often, the general public seem to view science and scientists only as a source of empirical knowledge (Quinn, 2009). They see science as knowledge only and not action. In fact, since the origin of our English word, “science”, is the Latin base *scire* which means, *to know* (Oxford Dictionaries, 2013), you might say that they take the word too literally.

Even when the process of science is acknowledged, it is commonly thought of as “The Scientific Method.” This notion is a vast oversimplification of the actual activity of scientists (Bentley, Ebert, & Ebert, 2007). The myth of a specific scientific method (Bauer, 1994) can be traced all the way back to the 17th century and the writings of Francis Bacon (Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002). Bacon and other empiricists held that there is a single reality that can be uncovered, that causality is always linear, and that objectivity is possible. As a result, empiricists believed that methodology guarantees unbiased results (Bentley et al., 2007). Here in the 21st century this error in thought remains alive and well in the minds of the general public. Most unfortunately, it also permeates far too many science textbooks and even the conceptual framework for some science educators (Bybee, 2002). Even though the views of positivism (often discussed interchangeably with empiricism) have long been discredited and replaced by postpositivism and constructivism (Bentley et al., 2007), the residual effects of its long-time stronghold can be seen. One reason for sure is that, though the early scientists’ views of inductive logic may be now seen as extreme, the method presented by Bacon, Newton, and others is easy to understand and therefore easy to accept (DeBoer, 1991).

If no one method for scientific practice exists, what then are the activities of scientists? Complexity in answering this question serves to explain the lasting presence of “The Scientific Method.” Scientific inquiry does indeed position itself on empirical evidence. Scientists collect data to explain natural phenomena, and the validity of this data is verified through checking measurements and repeating observations (National

Research Council, 2000). Within the confines of a typical classroom experiment contradictory data is regularly dismissed as student error—because what is being tested is accompanied by a canned answer. In opposition to this, practicing scientists are always on the lookout for inconsistencies. These are primary to the progress of science (Quinn, 2009).

Scientists gather data; however, much of the data and many of the methods of collecting data used by scientists are discipline specific. Scientists in chemistry and physics may conduct many controlled experimentation, yet those in the fields of astronomy or ecology must rely upon observation as a major source of data (Bentley et al., 2007). Scientists practice from the framework of accepted theory, and it is from this platform that hypotheses are described and tested. However, these practices are not merely to replicate what is already known. Instead, scientists explore hypotheses to further explain and develop current theory (National Research Council, 2007). In “The Scientific Method” the hypothesis is often an end unto its self—testing what is already known (Hodson, 2009). Reality shows that the actual process of science is messy and necessarily iterative. There are dead ends and inconclusive results, and even when answers can be found they inevitably lead to new questions (Quinn, 2009).

In *Taking Science to School*, the NRC (2007) presents the practices of scientists from a philosophical perspective by synthesizing key perspectives on the process of science. First, science can be seen as a process of logical reasoning. This view sees the scientist as a reasoned problem-solver—emphasizing the role of domain-general reasoning. The NRC provides evidence for this perspective by the likes of Inhelder and

Piaget (1958); Bruner, Goodnow, and Austin (1956); and Klahr (2000) . A second view is that science can be seen as a process of theory change. In this view, knowledge of science evolves as theories are elaborated upon with the support of new evidence—emphasizing domain-specific knowledge. Here, logic is not as important as conceptual change built upon evidence clashing with the theoretical framework. According to the NRC, this view is held by a number of theorists including Kuhn (1962); Chi (1992); and Carey (1991) . A final view on science is seeing science as a process of participation. This least cognitive approach views science as an activity or practice. From this perspective both logic and theory are elements of a broader and more inclusive practice of science. Such a practice includes both cultural and social influences. The NRC reports that researchers and theorists like Longino (2002); Thagard (1999); and Nersensian (2005) support this view of science.

Understanding that science is both an established collection of knowledge and a complex practice of discovery is essential since, “Your own concept of the nature of science is an important part of your content background and comes into play when you make decisions about what to teach and how to teach it” (Bentley et al., 2007, p. 7). Historically, science education in the United States has not successfully integrated these two essential elements of science.

There has always been a struggle over what is known to be best practices and the challenges that arise in the process of implementation (Science Education Curriculum, 2010). In science education, one place where this is clearly seen is in the desire for students to engage in scientific practice and difficulties in actually making it happen.

The importance of scientific practice in education has been understood since before the turn of the 20th century (DeBoer, 1991). Unfortunately, best practices in science education evidenced by research are rarely put into practice (Kulm, 2007). Reforms in science education have been no different. In the opening sentence of his forward for Bybee's (1993) book on reforming science education, Hurd said, "Since the early 1970s, there have been widespread efforts to change drastically the purposes for which science is taught. For the most part, reform has remained an illusion; the traditional mold has not been broken" (p. ix).

Nationwide Reforms in Science Education

The first nationwide reforms in science education took place beginning in the 1950s (National Research Council, 2007) and can be directly linked as effects of the success of the Soviet space program (Yager, 2000). These reforms made a sweeping decision about the debate over "science for scientists" versus "science for all." Internationally, this debate had been active since before World War I (Wellington, 2001).

In these first national reforms, led by natural scientists, science was presented as purely discipline based. Technologies and applications were removed; the curricula attempted to present scientists' knowledge and scientists' practices (Yager, 2000). This "authentic" science was an attempt to develop national appreciation for the value of science by creating a populace who understood the "science for scientists" (National Research Council, 2007). As such, it placed a high emphasis on the inquiry of science; however, did so through text-book driven instruction (National Research Council, 2007; Yager, 2000). The results of this period of reforms are not clear cut. Difficulties came

from every angle. Outside of instructional challenges and problems in implementation, parent groups and calls for local control followed quickly on the heels of the reforms. And while an attempt was made to emphasize both the knowledge and process of science, scientists and educators underestimated the challenge of the task (National Research Council, 2007).

The hope that national standards could succeed where so many other reforms have failed was stated clearly by Bybee (2006), “The power of national standards lies in their potential capacity to change the fundamental components of the education system at a scale that makes a difference” (p. 57). The need for and potential in national standards found solid footing in 1989 when the National Governors Association agreed to the idea of national standards. President Bush added his support, and educational organizations got to work (National Research Council, 1996). After a four year, seven million dollar effort, the NRC published the *National Science Education Standards* in 1996 (Yager, 2000). The goal of the standards was to produce a scientifically literate populace. Accomplishing this goal would involve four elements. (1) Students needed to not only know about the natural world but experience the joy in knowing. (2) Students needed to actually use scientific practice to make decisions. (3) Students needed to discuss and debate issues in science from a knowledgeable position. (4) Ultimately, students needed to become a productive part of society and their understanding of science should help to make this happen (National Research Council, 1996).

The thoroughness of the *National Science Education Standards* of 1996 was due in part to important and illuminating documents that had come before. These included

the AAAS's Project 2061 and its publication, *Science for All Americans* and the NSTA's publication, *The Content Core* (National Research Council, 1996). The publication of documents during the process (i.e. the *Discussion Document* and the *National Science Education Standards: A Sampler*) for the purpose of eliciting public comment (Collins, 1995) is a pattern that has been used again in the formation of the NGSS and serves to strengthen the final results (S. Metz, 2012).

Current Efforts to Revise National Science Education Standards

Even with the many admirably qualities of the 1996 *National Science Education Standards*, it is time to revise our national science standards. For one thing, some researchers and educators purport that improvements to science education in the United States are not all that impressive (National Research Council, 2007). Fortunately, there is evidence within current research for improvement. Rodger Bybee (2006), a key player in the publication of the 1996 Standards, lists two important revisions to be made, though he purports that they are minor changes. First, he says that the amount of content covered needs to be cut by up to 30% and the key concepts need to be clarified. Second, he states that the document needs to be made more functional for policy makers and curriculum designers.

The NRC (2007) argues for far more sweeping changes that go beyond a cosmetic update. They present eight recommendations for change and improvement in science education. The first three link directly to standards. (1) The framework of standards must change to meet what we now know about student thinking. (2) A few core ideas need to be identified and developed to demonstrate an increasing depth of understanding

through the years. (3) Standards must do a better job of presenting the process of science as a theory building and refining exercise and put students in the position to explore genuine questions.

Several documents recently published demonstrate just how major these changes have been and will be. In *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Research Council, 2012), the council presented their plan for developing standards that would meet current thought on teaching and learning. In the summer of 2013 the *Next Generation Science Standards: For States, By States* (NGSS) (NGSS Lead States, 2013) were published in completion of this major undertaking. Now, the task is left to educators, curriculum leaders, school districts, and teacher preparation programs. As with the *Common Core State Standards* (CCSS), most states will most likely adopt the NGSS in the coming years. Will educators be ready to implement the NGSS? How will we know when they are successful? Can a survey instrument be developed that assists in understanding the state of teacher *understanding* and *readiness* to implement as well as the needs for change and improvement? Desimone and Le Floch (2004) suggest that survey instruments are a good fit for evaluating the depth or degree of the implementation of an intervention. The goal of this research was to see if this held true with regard to *readiness* for the implementation of science standards.

Statement of the Problem

With the arrival of the new framework for science education—as presented by *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core*

Ideas (National Research Council, 2012), major shifts in science education are underway. A definite roadblock to this evolution is the current lack of clarity on how successful elementary educators will be in implementing the new framework. With no way to measure the level of readiness for implementation, it is difficult for state agencies and school districts to plan for professional development. It leaves teacher preparation programs at a disadvantage in attempting to prepare pre-service educators for the challenges that lie ahead.

Since most states have yet to adopt the NGSS, a way to assess current inservice educators' *understanding* and *readiness* is needed. The NGSS Lead States (2013) developed the standards based on *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Research Council, 2012). Therefore, *understanding* and *readiness* for the NGSS is contingent upon *understanding* and *readiness* for the NRC report (referred to in this research as, The Framework).

The purpose of this study was to identify key constructs of The Framework for successful implementation and to develop an instrument—the *New Framework of Science Education Survey of Teacher Understanding & Readiness (NFSE-STUR)*—to measure inservice elementary teachers' *understanding* of and *readiness* for The Framework. The study sought to identify critical factors for implementing The Framework at the elementary level and to validate an instrument that could be used by educators to self-assess these factors.

Research Questions

This study examined The Framework and related documents to determine the key themes for a successful implementation of the standards. Using these elements, the *NFSE-STUR* was created and validated. Two questions were used to guide the research.

1. What themes representing the implementation of The Framework are identified by the science education literature?
2. What are the underlying dimensions of *NFSE-STUR* items written to assess elementary teachers' *understanding* of and *readiness* to implement The Framework?

Limitations and Delimitations

This study has several limitations and delimitations that must be addressed. First, as a self-reported survey, the *NFSE-STUR* is working with participant perceptions. No data will be collected through observation. As such, differential understanding is an issue—people responding to the survey may have varying levels of familiarity with The Framework. Varying familiarity may influence prospective participant choice in responding to the survey. People who choose not to respond may do so out of insecurities with The Framework. There are also delimitations that impact this study. The population of the study will be K-5 elementary teachers who are invited to participate based on being on certain email lists (some specifically connected to science education organizations) or through forwarded emails beyond the influence of the researcher. Another delimitation is that the expert review will be limited to five experts.

Definition of Terms

- The Framework is used through this document to refer to the new framework for science education as described in, *A New Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Research Council, 2012).
- *New Framework for Science Education: Survey of Teacher Understanding and Readiness (NFSE-STUR)* is the instrument this research studies and is designed to measure elementary teachers' *understanding* of and *readiness* to implement The Framework. Throughout this document various version of this instrument are described. The draft used during the pilot study (*Pilot Draft: NFSE-STUR*). The draft used during the validation study (*Validation Draft: NFSE-STUR*). Half of the resulting data assessing teacher *understanding* (*NFSE-STUR Understanding Scale*). The other half assessing teacher *readiness* (*NFSE-STUR Readiness Scale*).
- *New Framework for Science Education: Survey of Teacher Understanding (NFSE-STU)* is half of the study's finished product. This is the validated version of the new instrument measuring elementary teachers' *understanding* about The Framework.
- *New Framework for Science Education: Survey of Teacher Readiness (NFSE-STR)* is the other half of the study's finished product. This is the validated version of the new instrument measuring elementary teachers' *readiness* to implement The Framework.

- The *Next Generation Science Standards: For States, By States*, frequently referred to as the NGSS, are the latest national science standards (NGSS Lead States, 2013). These standards replace the *National Science Education Standards* (National Research Council, 1996).
- Elementary is used in various combinations to refer to teachers in grades K-5.
- Themes is a term used in this document to refer to the broad overarching ideas found to best represent unique elements of *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Research Council, 2012). These “Themes” were used as foundational categories in developing the initial item pool.
- Integration is the combining of discrete parts into a unified whole (Case, 1991). Various types of integration are discussed throughout this research. Each presents a more specific aspect that falls under Case’s overarching definition.
- Factor analysis is a statistical measure that can empirically explain the underlying constructs of a set of items (DeVellis, 2003).
- Exploratory analysis is used throughout this document to refer to the factor extraction method of principal component analysis (Bangert, 2009).
- Principal component analysis (PCA) is the most common extraction method for identifying factors in a set of instrument items (McCoach, Gable, & Madura, 2013).
- Confirmatory factor analysis (CFA) is used to statistically test a hypothesized factor structure. This is done when researchers have some knowledge of the

underlying latent variable structure and, based on knowledge of theory and empirical research, postulate relationships between observed measures and underlying factors a priori (Bryne, 2009).

- Behavioral learning theories are theories that view learning in terms of the frequency of conditioned behavior.
- Cognitive learning theories are theories that view learning in terms of mental structures and the processing of information (Schunk, 2012).

Significance of the Study

This study is timely. The NGSS have just recently been completed and states all across the country are considering their adoption. An instrument measuring the current state of preparedness based on The Framework could be of great value in identifying needs in a successful transition to the NGSS. Individual educators could use the survey to determine areas of personal growth as a science teacher. Districts and building administrators could plan for professional development based upon local needs. Professional developers could use the instrument to effectively structure trainings to meet the needs of inservice teachers. State agencies could administer the instrument state-wide as a needs assessment measure. Teacher education programs could use regional data to beef up syllabi and programs of study.

CHAPTER TWO

LITERATURE REVIEW

Introduction

In this review of the relevant literature, topics crucial to the research are fully explored. Elements of science education are reviewed first. These include the current state of elementary science, the history of science education, and the science and psychology of learning science. Next, the NGSS are reviewed—demonstrating the inseparable connection to The Framework. Then, after reviewing relevant aspects of integration, The Framework is reviewed in light of integration. This involves a presentation of the history and value of integration, as well as, those aspects of integration critical to this study. Finally, integration is tied to The Framework and connections between The Framework and the CCSS are explained. The review then shifts from a focus on science education to instrument design. This process is given a complete review before identifying the themes within The Framework.

The Current State of Elementary Science

In 1995, the National Center for Educational Statistics (NCES), in cooperation with the International Association for the Evaluation of Educational Achievement (IEA), began to administer the Trends in International Mathematics and Science Study (TIMSS). This study assesses students in grades 4 and 8 on an international scale. It has now been administered five times: 1995, 1999, 2003, 2007, 2011 (data was not collected for grade

4 in 1999). In the latest study, 2011, about 500,000 students from more than 60 countries took in the assessments (National Center of Educational Statistics, 2013). Results from these tests reveal that the United States rate of improvement remains in the middle of the pack (Hanushek, Peterson, & Woessmann, 2012). Such mediocrity represents a failure to meet goals set for in the reforms of the 1990s—one of which was to have students first in the world in math and science achievement by 2000 (Collins, 1995; National Research Council, 2007). While the Nations Report card, which reports on NAEP testing, has shown an improvement of a year's worth of learning over the past two decades, only half of that has been seen in the TIMSS (Peterson, 2013). National Assessment of Educational Progress (NAEP) is a national assessment of science, reading, and math administered to 4th and 8th grade students. Peterson describes two logical explanations for the discrepancy. The NAEP tests could be getting easier, or the TIMMS tests could be getting harder. Regardless of the factors, one key element cannot be disputed when examining the TIMMS data in isolation, "U.S. students are not closing the international achievement gap" (p. 5).

In 1995, U.S. 4th grade students performed above the international average on the TIMSS science assessment. Only Korea out-performed the U.S and five nations were not statistically different (U.S Department of Education, 1997). Our 8th grade students while performing above the international average, were out-performed by nine countries and not significantly different than sixteen others (U.S. Department of Education, 1995). In the most recent study, 2011, U.S. 4th grade students still remained in the top ten with six nations out-performing and three not statistically different. Our 8th grade students are

still not performing as well. While remaining above the international average, U.S. 8th grade students were out-performed by twelve countries and not statistically different from ten others (Provasnik et al., 2012).

Caution is appropriate in drawing conclusions from the TIMSS alone. Tienken (2013) points out that the TIMSS creators readily acknowledge a mismatch in the assessments and some countries curricular organization. This includes assessing U.S. 8th graders on content they will not be taught until 9th or 10th grade. Another issue is with the most common aspect of the TIMSS to be reported and examined—nation rank (Nixon & Barth, 2014). This is true of previous paragraphs in this study. Nixon and Barth (2014) explain that merely examining national ranking fails to account for the cognitive domains for the questions and may result in invalid conclusions regardless of the validity of the instrument. Regardless of the cautions, these measures are used as part of the NRC's (2012) case for the need of a new framework in science education.

The History of Science Education

Understanding the evolution of science education requires a detailed examination of its history. This next portion of the literature review will examine the history of science education by considering a number of characteristics influencing each historical circa. Science education will be described in each circa by both the theoretical ideas as well as the actual practices of teaching science contemporary to the period. This will require a consideration of the nature of science, current learning theory, and the presence or absence of national reform.

Origins of Science Education (Before Circa 1880)

Modern science saw its beginnings in the 17th century and was solidly built upon the principles of empiricism as expressed in Francis Bacon's *Novum Organum* (Harris, 1975; Lederman et al., 2002). At the time, science was seen as a process of uncovering truth (National Research Council, 2007). This was the origin of "The Scientific Method" and the notion that a set of inductive steps guarantee results that can be purported as truth (Lederman et al., 2002). This empirical method relied solely upon direct observation of evidence through experimentation. Bacon, "protested against the prejudices of the learned and of the vulgar, against the fruitless debate without systematic investigation . . ." (Harris, 1975, p. 155). Such a narrow view of science is in direct opposition to a theory and model based science where scientists and the scientific community progress through conceptual understandings as much as empirical evidence (National Research Council, 2007). It suggests that current experimentation explores questions drawn from prior understanding but never constructs explanation from anything but factual data in the observation (Toulmin, 1982). Therefore, empiricism has no place for logical thought at all. The empirical view is that any "truth" obtained through logic is not science because it is devoid of observable evidence (Harris, 1975). This observable evidence positions the empiricist singularly dependent upon an experienced reality (Dilworth, 2006).

From this empirical foundation, science education merged with other forms of public education over a period of years. In the United States it appeared as early as 1750 with Benjamin Franklin's academy in Philadelphia. Over the next many years, the number of academies steadily grew. Because the academies focused on the practical

nature of life, they filled a role that many felt was missing in the narrow curriculum of classical education. It was in these academies that the study of science was first promoted (DeBoer, 1991); whereas, it found no place in the Latin grammar schools (Petersen, 1959). Thus, the academies—which would ultimately become our public high schools—offered an option for students who had completed their elementary, common school education but did not want to enter grammar schools en route to college (DeBoer, 1991). With science education only offered in the academies, science at the elementary level was nonexistent (Bybee, 1993).

Science education had found its way into the public school system.

Unfortunately, “textbooks were inadequate, teachers were poorly prepared to teach the wide range of subjects, and demonstration materials were in short supply and in poor repair” (DeBoer, 1991, p. 20). Plus, science courses were taught according to the pattern of education at the time—instruction consisted of students reading, memorizing, and reciting texts (DeBoer, 1991; Science Education Curriculum, 2010). With the first psychological laboratory not being founded until the end of this era (Schunk, 2012), the methods of instruction were influenced by both the empiricism of science and philosophy. With scientific empiricism demanding that knowledge be built on observation alone and philosophical empiricism presenting knowledge as being derived from sensation (Harris, 1975), it is not surprising that texts were to be taken as truth. While not equivalent, this thinking most closely aligns with conditioning learning theories—behaviorism (National Research Council, 1999; Schunk, 2012). Sensation, much like the behaviorist’s stimuli, was believed to be received passively (Harris, 1975).

As this era came to a close, the academy was replaced by the high school (DeBoer, 1991; Petersen, 1959) and science education made its first inroads into the elementary classroom (Bybee, 1993).

Early Structuring (Circa 1880 to 1920)

In a more accurate application of scientific empiricism, it was during this second era that science began to be defined as both a body of knowledge and a practice uncovering that knowledge. Operating from this perspective, educators and organizations began promoting a science education that included both knowledge and practice. When the National Education Association's Committee of Ten met in the late 19th century, they spoke strongly for a science curriculum emphasizing direct contact with the physical world. This would allow students to "be scientists" (DeBoer, 1991). Dewey (1976), as early as 1901, spoke in favor of an experimentation-based science education. He emphasized the importance of the body in learning about the world around us and referred to this as "manual training." The American Association for the Advancement of Science desired to see both scientific knowledge and scientific method dispensed to students (J. M. Cattell, 1914). A few years later the National Education Association (NEA) (1920) published their report on restructuring science education. In it they presented the importance of students learning science through their own experiences instead of by reading and hearing about scientists' experiences.

In light of this theoretical shift, it is somewhat surprising that the practice of science education showed little change from the era of its origin. It continued to emphasize only the knowledge of science at the neglect of the practice of science (Bybee,

1993). Though the predominate psychology of the time, functionalism, may have had a positive influence over the science education occurring in the schools, it never occurred. John Dewey promoted the idea, but because of the nature of functionalism it was impractical. “Although this goal was laudable, it also was problematic because the research agenda of functionalism was too broad to offer a clear focus. This paved the way for the rise of behaviorism as the dominant force in U.S. psychology” (Schunk, 2012, p. 10). Therefore, it is the influence of philosophical empiricism and its psychology cousin, behaviorism, which can be seen during this era of structuring and restructuring.

Thorndike (1914) opened his discussion of educational psychology by admitting the great complexity to human learning and recognizing the value in looking first at the simplicity of the lower animal mind. This is a pattern seen throughout behavioral psychology. Thorndike had his cats, Watson (1917) his rats, and Skinner (1950) his pigeons. And while Thorndike (1914) associated only the first of his four types of learning directly to animal behavior, Watson (1913) went much further in suggesting that, “The behaviorist, in his efforts to get a unitary scheme of animal response, recognizes no dividing line between man and brute” (p. 158).

The goal of the psychology of behaviorism was the prediction and control of behavior (Watson, 1913) which separated behaviorists from previous efforts in psychology—i.e. consciousness. Early behaviorists established themselves as the first scientific psychologists (Thorndike, 1914; Watson, 1913). They aligned themselves with empiricists in their approach to science—observing, deducing, and explaining. From this

they concluded that learning could be seen in terms of a connection between stimuli and response (National Research Council, 1999).

With behaviorism and empiricism both promoting observation and experimentation it is no surprise that this became more important in the schools. However, Dewey's experiential learning (Dewey, 1976) was not what these educators had in mind. The first elementary programs—Parker's unifying theme model, Jackman's nature study, Harris' relationship of ideas program, and Howe's Systematic Science Teaching—all had one thing in common. They all emphasized the accumulation of scientific knowledge. Most placed a premium on experimentation and observation; however, the goal was mastery of content (Bybee, 1993). Indeed, even the high school laboratory experiments that became important for entrance to some universities around the turn of the century placed an emphasis on the detail of the recordings, the neatness of the lettering, and the beauty of the images instead of the experiment explored (Petersen, 1959). The few laboratory methods promoting genuine student discovery or understanding over procedure; going beyond knowledge was seen as too time consuming (DeBoer, 1991).

Dewey was opposed to the idea that mental processes could be broken into discrete elements of stimulus and response (Schunk, 2012). This indeed was a failure of behaviorism as it held tightly to observable conditioned behavior. It could not account for understanding or reasoning (National Research Council, 1999). Even behaviorists admitted that the simplicity of this description could not account for the very evident complexity of human learning. So, early behaviorists added detail by elaborating on the

complexity of situations, simple versus series associations, and the connections of learning (Thorndike, 1914). After his bold statement, previously quoted, Watson (1913), admitted that the plethora of studies into animal behavior produced little impactful change in psychologists' understanding into the complexity of human learning. He continued by explaining that researchers cannot study a behavior in light of consciousness; therefore, even when learning is demonstrated through the observation of the behavior, much remains unexplained, "the effect of past habit upon present response, the range of stimuli to which it ordinarily responds, the widened range to which it can respond under experimental conditions . . ." (p. 160).

Restructuring Efforts (Circa 1920 to 1950)

During this era progressive theorists and educators found a stronger voice and had some success influencing actual teaching practice. From a theoretical position, the progressive message in opposition to traditional science education was prominent. The NEA (1920) described science education as typically involving the presentation of essential vocabulary for the memorization of definitions which were then tested for accuracy. The Committee on Science, part of the Commission on the Reorganization of Secondary Education, shifted the purpose for science education from the development of intellect to a holistic development of individuals as contributing and happy citizens. The committee recommended the use of themes as broad umbrellas facilitating the newly important understanding of science. Under these umbrellas, project and problem based methods should be used to organize content and facilitate student development of science

skills. This type of experience would ensure the easiest connection to the practical everyday world—another new emphasis (DeBoer, 1991).

However, with behaviorism as the prominent psychology and empirical philosophy still clinging to the world of science, the broad sweeping changes called for by progressives did not become the reality they envisioned. This was due in part to misinterpretations of their vision by textbook publishers, by outspoken opponents, and by the growth of testing which most easily assessed factual knowledge (DeBoer, 1991). Regardless of theoretical shifts, science as the accumulation of knowledge continued to be the standard in the schools (Bybee, 1993).

In the late 1930s Dewey wrote in opposition to, the still prevalent, traditional education. The position of traditional education, according to Dewey (1938/1988), was that disciplines consisted of bodies of knowledge and skill that must be delivered to the minds of students. The National Society for the Study of Education, writing around the same time, echoed Dewey's concerns of textbook dependent education (DeBoer, 1991). During these years even demonstrations supporting the content, much of it outdated, were conducted by teachers with students in passive observation (Tannenbaum, Stillman, & Piltz, 1965). Since the content of the texts used was outdated, it held little interest for students. Brown (1940) saw a plethora of real-world applications of science that were being neglected by these archaic methods.

Learning for understanding and application to life, demanded a learning theory that could embrace and seek to explain complexities beyond observable behavior. Some behaviorists, viewing their position with a small “b”, left the empirical tradition and

allowed for hypotheses exploring mental processes while maintaining methodological rigor. Others held to a radical capital “B” version of behaviorism (National Research Council, 1999). At this time, Skinner began his work in the psychology of behaviorism. In opposition to the broad complexities of functionalists (Schunk, 2012), the holistic, integrative notions of Gestalt (Humphrey, 1924), and the progressive demands for practical application, Skinner (1950) maintained that such extraneous notions need not be considered for learning to take place. True to behaviorism’s origins in empiricism, he dismissed the complexities of learning suggesting that learning theory is the realm for discussing such complexities. Importantly, Skinner separated empirical research from learning theory. He recognized the existence of mental complexity, but described it as undiscoverable, concluding, “That a theory generates research does not prove its value unless the research is valuable. Much useless experimentation results from theories, and much energy and skill are absorbed by them. Most theories are eventually overthrown, and the greater part of the associated research is discarded” (Skinner, 1950, p. 194).

This thinking continued to promote the ideals of a stimulus and response, knowledge-based curriculum. Therefore, even though the recognition of science education needing both content and practice was present from its initial days in the late 19th century (Science Education Curriculum, 2010) and experiential, practical learning was promoted by progressives throughout the first half of the 20th century (Bybee, 1993), in practice, science education remained solidly on the side of knowledge accumulation (DeBoer, 1991).

National Reforms and Elementary Science (Circa 1950 to 1980)

In the 1950s, elementary science education, as we know it today, was just coming into its own. Before this era, elementary science received little meaningful attention, “The science educational vacuum that has existed in the elementary school for many years . . .” (Ruchlis, 1964, p. 388). While there were sporadic efforts to bring science to the elementary classroom (as previously discussed), it was not until the early 1950s that it was considered an essential part of the curriculum (Vessel, 1963). Up to the 1940s most elementary schools had no science programs at all (Victor, 1965). And, even in the backdrop of the previously discussed history of the discipline, high school studies were still solidifying. For nearly 50 years the science course of choice had been “general science” (DeBoer, 1991). Petersen (1959) relays the results of a contemporary study into major U.S. high schools. Only some 22% were offering a unified physics course, and the unified biology course, common in the early 1900s, was just making a comeback.

From a theoretical position there was still strong support for balance in science education. Schwab (1962) stated the goal of science education as “enquiry into enquiry”, and in the many texts for science educators that begin to appear in the 1960s, there is a ubiquitous call for an active, experiment rich education that promotes both the knowledge and methods of science (i.e. Navarra & Zafforoni, 1960; Vessel, 1963; Victor, 1965). Though, not all were in agreement with this growing view of science education (Newton, 1968). World events interrupted and reinterpreted any debate about these ideals.

The Soviet success with Sputnik challenged U.S. supremacy and shook us politically, scientifically, and educationally. The result was national recognition in the

importance of science, the first national efforts in reform, and a national curriculum for science education (Bybee, 1993; DeBoer, 1991; Yager, 2000). These national reforms in curriculum included the involvement of practicing scientists in the development of a science education agenda (National Research Council, 2007).

This first national curriculum was developed by the National Science Foundation (NSF) through a process of twenty different projects and with the cooperation of natural scientists and psychologists. (National Research Council, 2007). The curriculum was built upon Bruner's model of science education (Bybee, 1993), but driven by the need of another generation of scientists, it was a curriculum of "science for scientists" (National Research Council, 2007). The idea was that it would be a true inquiry approach to science (Yager, 2000) and engage students in the actual practices of scientists. The key was the practice as it mirrored Skinner's operant conditioning of "learning by doing" (Schunk, 2012). The focus remained on scientific knowledge with scientific method playing a role as a means to the end (Bybee, 1993). Even though philosophers and scientists questioned scientific empiricism throughout the 20th century (DeBoer, 1991; Hawkins & Pea, 1987), these first reforms in science education were heavily influenced by this archaic view of science (National Research Council, 2007). Rigor was high and content was primarily determined by practicing scientists (National Research Council, 2007). The nature of science performed by scientists was seen as beyond the practicing educator and needed to be teacher proof (Yager, 2000). The curriculum attempted to educate science students separate from, even in spite of, teachers (Bybee, 1993). One outcome of this design was the modern resiliency of "The Scientific Method." It was

empiricism that propagated this notion, and is reflected in the science processes emphasized in the curriculum of the 1950s and 1960s (DeBoer, 1991).

The curriculum had the lofty goal of creating a student populace who appreciated and understood scientific practice (Yager, 2000); therefore, at the elementary level there was a strong emphasis placed on “process skills” (National Research Council, 2007). Two elementary programs that developed from the movement were Conceptually Oriented Program in Elementary Science (COPES) and Science—A Process Approach (S-APA) (Bybee, 1993). Finally, the philosophy of scientific investigation was embraced. Yet, the organization remained behaviorally structured. It was a canned curriculum and failed to transform the teaching of science. Even the labs were taught through direct instruction and tightly directed by textbooks. This produced a curriculum that was “teacher-proof” (Yager, 2000, p. 51). As a result, elementary teachers who lacked the needed training, failed to connect the knowledge in the text to the world just outside the classroom window (Klimas, 1969).

The teacher’s role as a co-developer of curriculum and a dynamic partner in the learning process had been neglected. Instead, “The curriculum framed discrete actions for teachers and students to create interactive science classes” (National Research Council, 2007, p. 14). In describing the NSF’s curriculum of the mid-century, Yager (2000) said, “However, even the inquiry approach, with more open-ended laboratories, tended to be taught by direct instruction—still prescribed and directed by textbooks. Textbooks, in fact, were not questioned” (p. 51). In this way, teachers and students waded through facts established by the scientific community. They did not participate in

the practice of science; they did not assist in scientific progress. In short, they did not perform true scientific inquiry.

An unchallengeable text accompanied by prescribed actions for both teacher and student, demonstrate the durable nature of empiricism and the predominance of behaviorism. During these years, Skinner made great advances in his research and refined behavioral learning theory (Collins, 2002), or the science of behavior as he preferred to think of it (Skinner, 1975). He recognized the complexities of human learning, but stripped them away by focusing on the behavior. Suggesting that regardless of what the process of learning actual was, the behaviors accompanying it would appear in the observable data (Skinner, 1950). This singular focus allowed Skinner to advance and refine behaviorism's backbone, stimulus and response, and provide detailed insights on operate conditioning (Schunk, 2012)—demonstrating his ability to use stimuli at various frequencies and during various tasks and subtasks to shape behavior (Collins, 2002).

The reforms of the 1950s and 1960s had lofty goals and met some of them: high levels of rigor, presentation of distinct disciplines, and students following prescribed structures of scientific practice. However, pedagogical neglect and isolation from practical application ensured the failure of the reforms (DeBoer, 1991). As a result, the major thrust of this reform effort had faded into its slot in history by the mid-1970s (Yager, 2000). What had been a time of public support and enthusiasm for science education shifted to criticism (Bybee, 1993).

Though its influence over education would be slow in coming, the late 1950s saw the inception of cognitive science and with it the possibility of explaining the complexities of human learning (National Research Council, 1999). Beginning with the work of Piaget in the 1960s, cognitive theories on development and learning rapidly rose in acceptance (Schunk, 2012). Instead of being confounded by the complexity of learning like early behaviorists (i.e. Thorndike, 1914; Watson, 1913) or dismissing it as less than science (Skinner, 1950, 1975), cognitive scientists embraced it. From its origins, cognitive science used a multidisciplinary approach to learning and studying learning (National Research Council, 1999). Such an approach to research allowed investigators to study holistically instead of departmentally—i.e. observable behavior. The result was that all twelve of Norman’s (1980) issues, “Belief systems, Consciousness, Development, Emotion, Interaction, Language, Learning, Memory, Perception, Performance, Skill, Thought” (p. 14), could be considered, even integrated in research.

National Science Standards (1980 to 2000)

After the failures of previous reforms, policy makers were determined to do better, and they got their chance a few years later. “In 1983 a new national crisis engulfed the political and educational establishments: the economic instability caused by the perceived supremacy of Japan, Germany, and other industrial nations over the United States” (Yager, 2000, p. 52). In light of this new threat, The National Commission of Excellence in Education delivered their report, *A Nation at Risk* (National Research Council, 2007). The report itself did little to produce successful reform. Nearly a

thousand mandates were issued requiring more courses, more rigor, more teacher-training, and a longer school day and year. Although very little meaningful change occurred (Bybee, 1993), *A Nation at Risk* was one piece of the puzzle.

Beginning in the 1970s, and gathering steam through the 1980s, the movement leading to national standards was marked by scientific literacy, scientific inquiry, and process skills (DeBoer, 1991). This era of science education saw a plethora of theoretical articles and empirical studies on the topic of inquiry. A broad ERIC search using the descriptors “science inquiry” yields over 1000 peer-reviewed journal articles since 1985. These efforts had the goal of developing both a scientific mind and an appreciation for true scientific inquiry (DeBoer, 1991).

Radical change was also underway in views of the nature of science. Empiricism was rapidly being replaced with postmodern science. Postmodern science had not completely shed elements of empiricism; it is after all an empirical practice. Observation and experimentation are important elements to true scientific practice; however, these scientists were willing to see the natural world with more flexibility. They were far more ready to turn to interpretation over arms-length explanation (Toulmin, 1982). By this time there were impossible limitations of empiricism in contemporary scientific thought. Science could no longer be described as simply a back-and-forth interaction of induction and deduction (Harris, 1975). It involved both conjecture and theory. It was best described as a constructed knowledge of a social population instead of a description of perceived reality based on certain procedures (Hawkins & Pea, 1987). And neither could these experiments and observations be random standalone enterprises, but rather built

necessarily upon prior theory. In order for such theory to exist, causality must be applied. A completely empirical deduction did not allow for this (Harris, 1975). It was an important part of scientific methods, but only a part.

In its most basic application, empiricism called for the scientist to observe the natural world without interfering. Not that this purist notion has ever held true in practice. Science has always been an interactive practice. Regardless, scientists now more fully embraced this reality (Toulmin, 1982). Also, instead of viewing scientific knowledge as a deductive generalization, science was now seen as interpretive (Hawkins & Pea, 1987).

By this time cognitive science had greatly grown in strength and produced an abundance of related learning theories. These quickly took a prominent place in educational psychology, if not in every classroom practice. The success of cognitive science can be attributed a couple key elements: technological and methodological advancement and multidisciplinary perspective (National Research Council, 1999). Common to all the cognitive theories was the assumption that learning is an active process within the brain and that the researcher's attention needs to focus on this cognitive activity instead of the resulting behaviors (Collins, 2002). This study of the "process of knowing" is a key attribute in cognitive science (National Research Council, 1999) and the generally understood process is that students construct knowledge (Mestre & Cocking, 2002) by assimilating new knowledge into current brain structures or by accommodating new information by reorganized existing structures (Piaget & Inhelder, 1969) from whence comes the term, constructivism which, in general terms, states that

individuals construct their own knowledge as they interact with ideas, their environment, and others (Glaser, 1979; Moshman, 1982; Piaget & Inhelder, 1969; Schunk, 2012).

Constructivism is not technically a theory of learning but rather a psychological and philosophical perspective; it is an epistemology (Schunk, 2012). Under its broad umbrella, there are diverse viewpoints. These can be categorized as exogenous, endogenous, and dialectical constructivism (Moshman, 1982). It is the later viewpoint, dialectical constructivism, with which the majority of the constructivist learning theories most closely align—i.e. social cognitive theory, motivational theories, Bruner’s development theory, and Vygotsky’s constructivism (Schunk, 2012).

By the beginning of the 1980s many cognitive theories had gone through a number of years of development and testing; however, much remained a mystery and had little practical application, “Perhaps one reason that existing theories say so little relevant to real world activities is the neglect of social and cultural factors, or emotion, and of the major points that distinguish an animate cognitive system from an artificial one . . .” (Norman, 1980, p. 1). Never the less, cognitive research was changing instruction. Piagetian theory, as interpreted and applied to education through the likes of Bruner (1960) and Lawson (1979) (DeBoer, 1991), had a great influence on theories of teaching. These thoughts on human development changed instructional strategies models for teaching to be appropriate with concrete or formal operational stages (Linn, 1982).

While some teachers readily mentally embraced these new ideas, classroom practice changed little (Bybee, 1993). Educators tend to teach with strategies that have been used for years—not because they have established them in light of research but

because they are “tried and true” (Pratt, 2002). Therefore, what had been done for years in light of the psychology of behaviorism was regularly practiced during this era. In fact, much of what occurred in classrooms was dictated or at least influenced by the legacy of behaviorism from learning objectives to systems of management (Duschl, 1990).

Thorndike’s (1914) century old work on habit forming was still used during these years in the drill and practice instructional strategy (Collins, 2002). And while Collins suggests that such a notion is fine in the development of habits like putting on safety goggles when Bunsen burners come out, the reality is that such principles of behaviorism have influenced curriculum and instruction in far deeper ways. “Textbooks are filled with facts that students are expected to memorize, and most tests assess students’ abilities to remember the facts” (National Research Council, 1999, p. 9). Even attempts to produce a more active, engaging pattern for learning science, often fell into behavioral traps.

As this era built to its crowning achievement, the *National Science Education Standards* (National Research Council, 1996), it was recognized that even with the many successes, science education had far to go. And there was still much debate about where the answers lay,

We are in the midst of a great educational reform, and one of the most pressing questions has to be: “What should we do?” Everywhere there are recommendations, each set with the unique perspective of the group presenting the report. Yet we continue asking the same questions and seeking answers to the questions of what we should do and what we need. Perhaps we need national standards for science education. Perhaps we should improve assessment or promote new science curricula. Perhaps the answer is teacher education. (Bybee, 1993, p. 147).

In answering his own questions, Bybee (1993) pointed to a diversified leadership in facilitating change in all of the areas: national standards, curricula, and teacher education.

Then, he closed the discussion with a reminder of previous failures. Multiple attempts at reform since the Second World War had significantly changed policies and influenced programs; however, it had not changed classroom practice. “The rhetoric and the reality of reform do not conform. If we do not confront this issue, the contemporary reform will be recorded only as one of reports and recommendations, with no response” (Bybee, 1993, p. 170).

Building from *A Nation at Risk* (National Commission on Excellence in Education, 1983), *Science for all Americans* (American Association for the Advancement of Science, 1990), and *Benchmarks for Science Literacy* (American Association for the Advancement of Science, 1993), the NRC developed the *National Science Education Standards* (1996), and set forth goals in science education. These included students being excited about understanding the natural world, applying scientific process in their lives, and intellectually debating scientific matters (National Research Council, 1996).

Science Education in the 21st Century (2000 to 2013)

With the arrival of the era of science standards, “Science for all” was back and never more strongly emphasized. A science education that would balance both knowledge and skills of inquiry was the goal (American Association for the Advancement of Science, 1990; National Research Council, 1996). Unfortunately, data from the TIMMS over the past two decades (presented in detail in a previous section, The Current State of Elementary Science) indicates that even this most concerted and systematic national effort has failed to produce the results desired. The NRC (2007) summed up the lack of progress.

After 15 years of focused standards-based reform, improvements in U.S. science education are modest at best. International comparisons show that many U.S. students fare poorly relative to their peers in other countries. In addition, large achievement gaps between majority students and both economically disadvantaged and non-Asian minority students persist in all school subjects, and they are especially strong and persistent in science. (p. 11)

It seems that Bybee's (1993) warning has come true. Policy was changed radically, programs have been updated, but instruction remains little impacted. Even after all these years, science education has not shaken free from this notion of fact over theory and explanation over interpretation. Many textbooks and curricular materials still present an accumulation of "facts" (Bentley et al., 2007). Fifty years ago, Atkin's and Karplus' (1962) learning cycle described a specific set of behaviors for both teachers and students; it ignored what was occurring in the mind and assumed that by following the pattern learning would just occur (National Research Council, 2007). The learning that would result from such situations would be mostly factual. Today more than ever, understanding is prized over knowing. Yet, classroom practice is regularly aligned with the old model. Students are rarely given the time it takes to develop understanding, instead curricula and lesson objectives emphasize the accumulation of knowledge (National Research Council, 1999).

The NRC (2012) boldly claim that it is time to align science education with scientific practice which itself has undergone a paradigm shift. Increasingly, science is investigated by groups of scientists, where individual's ideas readily influence the thoughts and interpretations of others. The primary views on the process of science now include, (1) "Science as a process of logical reasoning about evidence", (2) "Science as a process of theory change", and (3) "Science as a process of participation in the culture of

scientific practices” (National Research Council, 2007, pp. 28-29). These are all a departure from empiricism. Toulmin (1982) argues that while various purists have suggested a return to empiricism, this notion has always been deceptive. Researchers may recognize rational, conceptual, and interpretive elements in science more readily now, but they have always been there. The implications, for science education, from these and other elements of the current view of the nature of science are addressed in the next section of the chapter, “The Science and Psychology of Learning Science”.

The failure to achieve lofty goals established by the reforms of the 1990s was not only associated with teachers’ reticence to update teaching practice, there are also fundamental issues that must be applied to the development of national standards for progress to be made. New standards must reflect our most recent understandings of human cognition both in their scope and structure. At the time, Skinner’s (1975) objection to the need to look “inside” for answers to learning carried much practical weight. He suggested that for mankind this is a natural urge, “We almost instinctively look inside a system to see how it works. We do this with clocks, as with living systems” (p. 44). He argued that we would not, could not, find the answers to learning by looking inside. Technological and methodological advancement has nullified his objection. Today technology can show the brain functioning during various mental operations (Schunk, 2012) which allows neuroscientists, working in concert with educational psychologists, to interpret its activity with increasing clarity (i.e. Dux et al., 2009; Pashler, 1994; Sirois et al., 2008). The NRC (2007) points to progress made in tools and methods (i.e. Anderson, 1987; Ericsson & Charness, 1994) to explain how the landscape

has changed. Implications of these recent capabilities are addressed in the following sections, “The Science and Psychology of Learning Science” and “*Next Generation Science Standards’ Connection to The Framework*”.

Throughout the history of science education a number of factors have influenced theories on teaching science as well as actual classroom practice. Contemporary views on the nature of science often shaped the curriculum being taught and the methods employed. Philosophy and psychology also shaped science education encouraging certain instructional practices over others. The presence or absence of national reforms probably has had the most sweeping influences. Yet, of interest is the ubiquitous discrepancy between the theoretical ideals in science education and actual classroom practices. From the earliest days of public school science, policy makers, theorists, and educational researchers have promoted a science education that balances both the content of science and the practice of science. To date, this lofty notion has not seen popularity in the classroom. Added to this issue, theorists and policy makers have at times been behind the curve themselves. Texts of the 1930s and 1940s were often outdated. The reforms of the 1950s and 1960s dripped with the influence of empiricism—even though accepted understandings of the nature of science were beginning to shift away from its empirical foundations. Educators have regularly failed to embrace current practice either from ignorance or convenience. In the past, science inquiry was seen as too time consuming and life application excused as trivial.

Now, over the last 30 years, the value of genuine inquiry has been promoted by contemporary views on the nature of science, accepted understanding of learning through

the work of cognitive scientists, national reforms and policies, and teacher education. Yet, predominately, classroom practice remains fact oriented, behaviorally structured, and an exercise in empiricism. In an attempt to finally breakthrough all of these issues, the National Research Council (1999, 2007, 2012) has coordinated a massive effort to completely reshape science education.

The Science and Psychology of Learning Science

Isaac Newton famously said, “If I have seen further it is by standing on the shoulders of giants.” Newton puts into words an admission that all must make. It is no different for those attempting to orchestrate the next great step in science education. Current understandings of science and psychology have been built from the labors of those gone before. Today policy makers are informed by past successes and failures. Current understandings about how elementary students learn science is an integrated blend of ontology and epistemology. This portion of the review explores these understandings through a lens provided by the NRC’s (2007) report *Taking Science to School: Learning and Teaching Science in Grades K-8*. In this landmark work, the NRC presents four strands for scientific proficiency.

1. know, use, and interpret scientific explanations of the natural world;
2. generate and evaluate scientific evidence and explanations;
3. understand the nature and development of scientific knowledge; and
4. participate productively in scientific practices and discourse. (p. 36)

How Children Know and Understand

Over the past several decades research has significantly changed our understanding about how children know and understand. These changes hold important

implications for science education. First, diligent study has shown that viewing children as concrete and simplistic thinkers is erroneous. Since this is a commonly held notion, current practice needs an update. Second, vast research has been dedicated to prior knowledge and conceptual change and has demonstrated flaws in current curriculum and instruction.

Piaget is often credited with the origination of the stages of child development; however, he himself rejected such an association by separating his work from that of child psychology (Piaget, 1967). The “credit” more accurately goes to the work of Bruner, Lawson, and Karplus who took Piaget’s work and mishandled it. Piaget himself viewed and presented his work as a genetic epistemological theory and a psychologist dealing in children’s structure of knowledge (Elkind, 1967; K. E. Metz, 1995). Through this loose interpretation of Piagetian thought, the legacy of cognitive development as a series of distinct stages took hold (National Research Council, 2007). Research has demonstrated that such a neat and tidy explanation is not seen in reality.

For example, according to outdated thinking, young children should be given lots of opportunity to organize and categorize to develop their concrete thinking; however, children can think abstractly and participate in deep scientific inquiry beyond “hands-on” activity (K. E. Metz, 1995). When categorizing, children also develop theoretical ideas about the patterns and attempt to understand through inferences (Gelman & Markman, 1986). There is little doubt that the conceptualization and inferences of a young child differ from a ten year old or adult (Carey, 1991), yet the NRC (2007) concludes that if children did indeed enter elementary school as “concrete, preoperational, precausal,

prelogical, and lacking the ability to think in relational terms . . . [they would] bring a radically different way of understanding the world with them” (p. 55). Situated cognition, a core foundation of constructivism, may explain the discrepancy more accurately. Situated cognition explains that mental process is always located in physical and social contexts (Schunk, 2012). This employs a complementary view of both Piagetian constructivism and Vygotskian sociocultural theory to describe how learning occurs (Cobb, 1994). Instead of seeing younger students as developmentally unable for certain levels of thought, it could be helpful to explain their limited sophistication in terms of exposure to the needed physical and social contexts.

It may go without saying that children enter school with a rich conceptualization of the world around them, but the implications of this reality have often been misunderstood or ignored. This conceptual framework is a rich and essential starting point for learning, yet its misconceptions can be stubbornly resistant to change (i.e. Vosniadou & Brewer, 1994) (Duschl, 1990; National Research Council, 2007). In attempting to understand these implications, both cognitive scientists and educational psychologists have researched conceptual change—which is closely associated with information processing theory (National Research Council, 1999).

Long outdated views on learning emphasized the stimulus of imparted knowledge to produce behavioral change (i.e. Skinner, 1975). More modern ones emphasized inquiry and discovery skills as an avenue to the growth of brain structures (Screen, 1986). However, recent research has produced the surprising reality that, regardless of learning theory, new knowledge and understanding is not sufficient to produce a conceptual

change (National Research Council, 2007). The first fails to account for students bringing prior knowledge to the current learning (Yager, 2000); the second fails to appreciate the differences between a naïve child conceptualization and that of a fully developed adult (Carey, 1991; National Research Council, 2007). Time will not ensure that students replace their initial conceptual understandings with sophisticated scientific ones. This takes careful and intentional strategies on the part of educators (Duschl, 1990).

While conceptual change is often challenging, it is that prior, often naïve, knowledge that becomes the foundation for growth (National Research Council, 2007; T. D. Zimmerman & Stage, 2008). Knowledge is constructed by assimilation or reconstruction (Piaget & Inhelder, 1969). It is not done through deletion. Conceptual change occurs during Piaget's reconstruction (Carey, 1999) thus requiring prior understandings to develop. From this beginning understanding, students need to make connections to the new situation and the new information through various self-directed and teacher-directed practices in order for conceptual change to occur (National Research Council, 2007).

Nersessian (1989) argues that these practices should mirror scientific discovery which is constructed through problem-solving strategies and evaluative heuristic procedures. These include analogy as a way to make sense of new information or even to practice metacognition, thought experiments—regularly used by scientists, and imaginistic reasoning (Beeth, 1998; Carey, 1999; Nersessian, 1989). These practices allow students to construct new conceptual models through extension, combination and

modification (National Research Council, 2007). The NRC also purports that symbolic tools are key in concept development and conceptual change: written and spoken word, images and diagrams, and mathematical representations. The ideas most closely align with the social mediation of tools and symbols prevalent in Vygotskian constructivism (Schunk, 2012).

How Children Generate and Evaluate

Recent research into students' scientific reasoning has uncovered some encouraging and actionable results. Contrary to the historical view of young children as concrete and simplistic thinkers, it has been demonstrated that students are competent in their reasoning; whereas, adults are less sophisticated than might be assumed. Unlike theories in age-bound stages of development, the range in complexity of scientific reasoning is tied to conceptual understanding of the phenomena being studied (National Research Council, 2007). These realities have far-reaching implications for science education. Students of all ages need to participate in the generation of knowledge by asking questions, developing models, investigating, analyzing data, using mathematics, and constructing explanations. And, evaluate current knowledge by engaging in argument and evaluating information (National Research Council, 2012). There is no expectation that student participation will occur with equal sophistication; however, their involvement develops reasoning that can reach beyond their age (National Research Council, 2007). Sophistication beyond expected levels further demonstrates the importance of situated cognition.

How Children Understand the Construction of Scientific Knowledge

A key shift in science education becomes evident when researchers compare students' understanding of the nature of science with that of the scientific community. The mythical *Scientific Method* remains a primary understanding of science (National Research Council, 2007). Part of the problem is that science texts often present overly simplified views of scientists practicing in this neat and tidy fashion (Carey & Smith, 1993). Kuhn and Dean (2005) also point out that educational psychologists have further influenced this by emphasizing the development of scientific reasoning through the manipulation of variables. Students then see science as an empirical set of facts—they might “rediscover” through experimentation—that are not questioned, instead of the rich reality of the theory-building process (Carey & Smith, 1993; National Research Council, 2007, 2012; Rubba, 1981). While theorizing can be argued to be the defining feature of scientific thought (Hodson, 2009), the empirical view of factual knowledge that trusts the sensory experience alone pervades students' understandings into their college years (National Research Council, 2007). Instead, “We want [students] to come to understand that our knowledge of regularity in nature is a consequence of successful conjecture, rather than its precursor, and that an adequate theoretical perspective is essential to both observation and experimentation” (Carey & Smith, 1993, p. 236).

At the expense of an appropriate view of theory, explanation, and models, children view science empirically and historical science education has further entrenched this misunderstanding. Fortunately, research shows that a change in science curricula and instructional practice can remedy this distortion (National Research Council, 2007).

Demonstrating the complexity of science instruction, research shows that this change in understanding is facilitated more rapidly by explicit rather than inquiry-based teaching (Khishfe & Abd-El-Khalick, 2002). Teaching the history of science has also shown some potential in changing student view on the construction of science knowledge (Solomon, Duveen, Scot, & McCarthy, 1992). Plus, researchers have identified important curricular features like deep science problems, opportunities for genuine inquiry, and explicit discussions; however, it is important to note that more research into the development of student understandings of science is needed (National Research Council, 2007).

How Children Participate in Science

If students are to successfully participate in science, teacher instruction needs to be modified for their success. Research demonstrates that the norms in evaluating, explaining, and discussing scientific evidence are different from everyday life (Michaels, Shouse, & Schweingruber, 2008). Also, if motivation is lacking, students fail to fully engage in the discourse which hampers their ability to understand science at deep levels. Fortunately, there are things that educators can do to facilitate student participation and attitude.

Discussing science is extremely important in the process of advancing scientific knowledge, but it occurs quite differently from how students typically converse. A cornerstone activity of scientists is argumentation. They do so to persuade others that their ideas are correct (Bazerman, 1988). Bazerman explains that a study of Newton's science articles reveals excellent argumentation—writing constructed so that the reader will naturally and logically conclude that Newton is correct. Unfortunately, scientific

argumentation rarely occurs in the classroom. Filling its place is teacher directed talk attempting to convince students about something. The “turn-taking” school oriented pattern inhibits the occurrence of true argumentation (National Research Council, 2007). The NRC synthesizes the literature, demonstrating the crucial points for true argumentation to take place in classroom settings: teachers must become comfortable with argumentation and teach a “civil exchange”, the emphasis must remain on plausibility and evidence for the growth of collaborative understanding (unlike typical argumentation which holds the goal of “winning”), teachers must have adequate knowledge of the content and children in order to best facilitate, teachers and students must become comfortable with considering multiple explanations for phenomena, students must be given opportunities to work collaboratively with the facilitating support of the teacher, and teachers should consider assigning roles for group members to further encourage questioning and claim challenges.

Students who are less motivated will not fully participate in class. This motivation is influenced by the students’ values placed on science as well as their perception of their own ability to do science (National Research Council, 2007). Simpson and Troost (1982) purport that motivation in science education is influenced by attitudes towards science and a self-conception of science (as cited in Anderman & Young, 1994). Self-efficacy also modifies student motivation in science. Students with high efficacy are more motivated than those with low efficacy (Anderman & Young, 1994).

Engagement in scientific practice can be achieved—overcoming the challenges of motivation and attitude (National Research Council, 2007). Teachers have the ability to influence student motivation through classroom climate, curriculum (Simpson & S.J., 1990), and methods of instruction (Anderman & Young, 1994; Simpson & S.J., 1990)—specifically providing appropriately challenging activities, empowering students with control of their own learning while maintaining peer and teacher critique, and emphasizing scientific argument through evidence and logic (National Research Council, 2007). The NRC (2012) also recognizes the role of real-world application and subject-area integration in facilitating student motivation to learn.

Building from these four proficiencies, the NRC (2012) constructed a framework for practically employing their vision. With this framework as a foundation, The NGSS Lead States (2013) developed the new standards so that students participate in genuine investigations and solve genuine problems. Each standard of the NGSS demonstrates interdisciplinary connections through cross-cutting concepts and connections to the Common Core ELA and mathematics standards (National Research Council, 2012; NGSS Lead States, 2013).

Revisions to Teacher Education

Since the late 1950s and early 1960s, one word could capture theoretical understandings of science education all the way up to today. That word, “inquiry” (DeBoer, 1991). However, moving from theory to practice is a major undertaking that requires effective teacher training. As Bybee (1993) noted, classroom practice does not change merely because theoretical rhetoric has changed. In looking back, much of these

years were lost to practical change due to the lack of clarity over whether inquiry was an instructional method or a nature of science (DeBoer, 1991). The NRC (2007, 2012) has now abundantly answered those answers with a unified yes. Inquiry, now termed Science & Engineering Practices, is both a way of instruction for student learning and the nature of science. Teacher training then has grown closer to meeting best practice over the past half century, and it must continue to mature.

In the 1950s, science education at the elementary level was still in infancy. As such teachers already in practice had little to no training in teaching science. The result of this was predictable, “Reading, recitation, and writing. This is undoubtedly the most common method of teaching science. The teacher assumes that if the student has read the science book and can name the kinds of clouds, describe machines, or list major electrical inventions, he has learned science” (Vessel, 1963, p. 46). Teacher education programs struggled to respond to a need they were not prepared to fill. Empirical research to lead these colleges was scarce. The majority of guidance was theoretical and not research-based at all (Bruce, 1969). The lack of empirical foundations did not diminish the great efforts made to improve the training teachers received (Yager, Lunetta, & Tamir, 1979) or curriculum programs developed to assist them (D. J. Kuhn, 1973). Unfortunately, effort alone cannot lead to any sort of predictable success. Yager et al. (1979) reports on several large-scale studies of many education programs during these years and easily points out the lack of consistent approaches and the lack of a common goal. While most at least were attempting to head towards the idea of inquiry, dissenting opinions added confusion to the effort. Newton (1968) referred to inquiry based teaching as “dishonest.”

Claiming that it was unfounded, he continued to promote lecture and the memorization of a collection of facts. The ever lingering influences of behavioral learning theory also confounded inquiry efforts. Objectivism (the belief that knowledge exists of its own outside of the learner and can be experienced in an objective way) promoted science teacher education where knowledge would be “transferred” from master teachers who knew what was important to novices who did not know and would “accept” this knowledge (Tippins, Nichols, & Tobin, 1993).

Through the 1980s and into the 1990s researchers began examining, in earnest, what produced changes in teacher practice (Tippins et al., 1993). This produced more findings in ten years than could be compiled from the previous hundred (Yager, 1993). Clarity and uniformity in best practice began to build at the policy level. In 1991 the National Science Teachers Association set recommendations for science teacher education and exemplar institutions. They promoted twelve semester hours of science content, a hands-on science teaching methods course, field experiences with science education, qualified instructors engaged in research, and support of affective understandings in the importance of science education (Reiter & Peinich, 1991).

Unfortunately, little of this research was making its way into teacher training programs. In any event, embracing these ideals in practice was to prove a challenge since teachers and teacher preparation programs struggled to overcome the “teach as you were taught” model naturally found in our memories (Yager, 1993). Added to the practical difficulties were the practical realities, “Although psychology has moved away from behaviorism, education continues to embrace it. Teachers are being evaluated on the basis of

adherence to a set of procedures that . . . are inappropriate for [science education]”

(Howe, 1988, p. 310).

In spite of these challenges, constructivist principles were beginning to take shape in science teacher education. The individual and social construction of knowledge had been embraced by the education community (Tippins et al., 1993), and the recognition that changing science education began with changing teachers was being said by influential experts,

The success of the education reform movement depends on teachers who have the knowledge, disposition, and skills to teach in ways that reflect the goals of the new programs. In the long term, improvement of the science knowledge and teaching skills of teachers is the single most important environmental outcome to be derived from the development and implementation of the new programs (Bybee, 1993, p. 114).

As the decade of the 1990s advanced, teacher education programs had to negotiate strategies of addressing and best using the various policy documents emerging (i.e. *Benchmarks for Science Literacy* and the *National Science Education Standards*).

Looking at the past fifteen years of teacher education, there has been continued progress towards teacher training in the best practices of science education. Since an instructional model that is research-based is vital to long term success (National Research Council, 1999), it is encouraging that Bybee (2006) reported the 5E model for instruction appeared in over 97,000 posted examples in university course syllabi. During these years, the continued promotion of the nature of science (Lederman, 2011) and the essential elements of inquiry (National Research Council, 2000) has kept these elements at the forefront of research into the preparation of science teacher (Yager, 2005). In addition, advancements in the science of learning and the embrace of these principles

(National Research Council, 2007) have also produced new focuses including, better understanding of student learning, more effective assessment strategies, and developing an attitude of inquiry in teaching (Yager, 2005) During these years, the importance of teaching for conceptual change and using progressions of learning instead of the historic scope and sequence model also began to receive attention (Michaels et al., 2008).

Now, the community of science education finds itself in the midst of yet another transition. The term, “inquiry” has been set aside for the phrase, “Practices of Science & Engineering” (National Research Council, 2012), four proficiencies, instead of the three from the 1990s, frame the task of science teacher educators (National Research Council, 2007), and most states face the near adoption of the NGSS. Success in these efforts will not occur without teacher training. And the partnership of all players in the science education system, “[Which] includes organization and administration at state, district, and school levels as well as teacher education, certification requirements, curriculum and instructional resources, assessment policies and practices, and professional development programs” (National Research Council, 2012, p. 241).

Next Generation Science Standards’
Connection to The Framework

This section reviews the *Next Generation Science Standards: For States, By States* (NGSS) by demonstrating the standards inseparable tie to The Framework. The Framework was used to build the NGSS and is both ontologically and epistemologically different than previous attempts at national reform. It was purpose written for the development of new standards and involves three integrated dimensions: Scientific &

Engineering Practices, Crosscutting Concepts, and Disciplinary Core Ideas. Each of these elements is reviewed in turn.

The great effort that was the *National Science Education Standards* (National Research Council, 1996) has not succeeded in closing the international gap (Peterson, 2013). The pattern of reforms failing to produce significant change (Bybee, 1993; Kulm, 2007) has repeated itself once again. Current science education has failed to achieve the sweeping goals established nearly two decades ago. This is in part because of a lacking organization across grades, an overemphasis on a breadth of facts, and lacking engagement in the practices of science. New standards must reflect what is understood about student thinking, emphasize a few core ideas that could be developed to demonstrate an increasing depth of understanding through the years, do a better job of presenting the process of science as a theory building and refining exercise, and put students in the position to explore genuine questions (National Research Council, 2007).

A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas (National Research Council, 2012) was constructed to deal with these weaknesses. The NRC (2012) determined that there was significant room for improvement based on reform documents, progression of science, lessons learned from standards-based education, and the growing knowledge on learning and teaching science. Collectively these aspects required a new conceptual framework. There has been growing national concern over the fragmentation of topics and the lengthy lists of facts to be learned in science curricula and has led to the now infamous “a mile wide and an inch deep” eulogy (Bybee, 2006; Coleman & Zimba, 2008; Li, Klahr, & Siler, 2006; Sneider

& Workosky, 2009). In attempting provide a framework that unifies science education, the NRC (2012) operated from three guiding principles: learning is a developmental progression, core ideas in science and engineering must be limited and both within and across the disciplines, and learning about science involves the integration of knowledge and practices to participate in inquiry or design. Each of these elements and principles are evident throughout the NGSS both in content and design (NGSS Lead States, 2013).

Scientific & Engineering Practices

Balancing the knowledge of science content with the practices of science has always created tension. Traditionally, even when the need for scientific inquiry has been acknowledged, science education has leaned strongly toward content over practice (Bybee, 1993; DeBoer, 1991; Yager, 2000). Guarding against a swing too far the other direction, the NRC (2012) choose to “use the term ‘practices,’ instead of a term such as ‘skills,’ to stress that engaging in scientific inquiry requires coordination both of knowledge and skill simultaneously” (p. 41).

Including the practices of science and engineering was important in The Framework, and ultimately the NGSS, because it helps students understand how our knowledge of science develops, how engineers practice, and the connections between the disciplines. Practicing science and practicing engineering also aides in deepening student understanding of the crosscutting concepts and core disciplinary ideas that form the other structural elements of the framework (National Research Council, 2012).

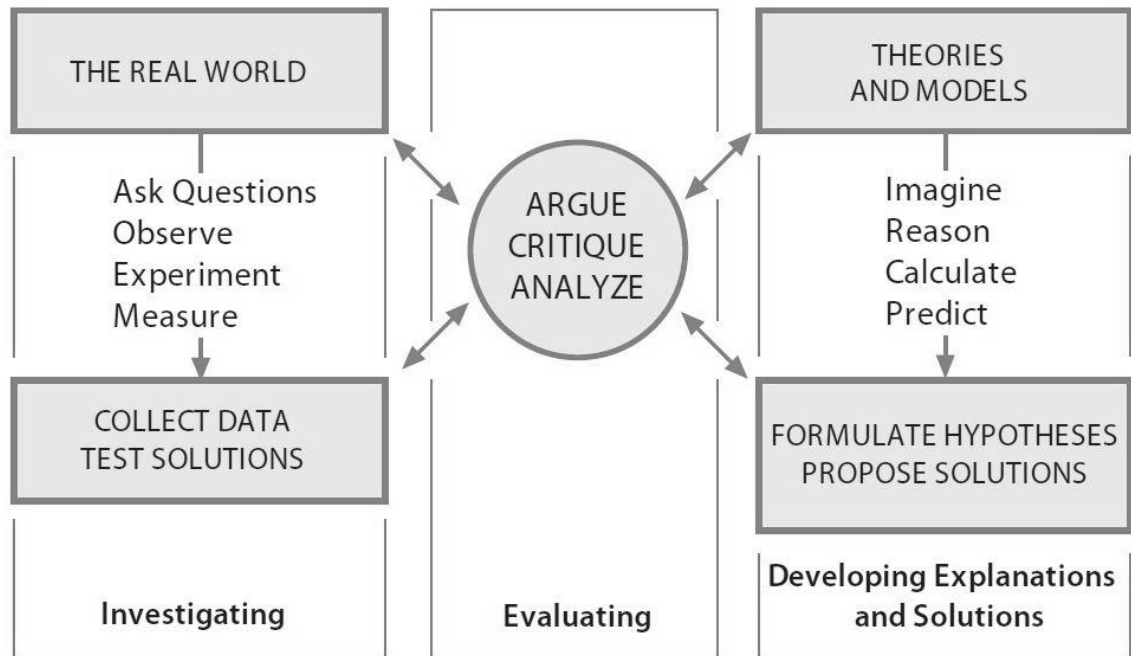


Figure 1. Unifying the Practices of Science & Engineering (National Research Council, 2012, p. 45)

While science and engineering do differ, their practices can be integrated into a unified model: investigating, evaluating, and developing explanations and solutions (See Figure 1). Unlike the idea of a sequential-step method, the figure is not intended to move neatly from left to right. All three spheres are iterative as argumentation and critique create the need for more investigation or demand new explanation. Practicing science and engineering in this way allows students to ask questions, define problems, use models, carry out investigations, interpret data, construct explanations, design solutions, argue from evidence, and communicate information (National Research Council, 2012).

Crosscutting Concepts

Crosscutting concepts are big ideas that have value across the disciplines of science and engineering. They provide a mental structure for students to use in

organizing new information. Including these concepts in The Framework gives teachers the explicit task of providing regular opportunities to connect knowledge and skill in this way. Previously, students have often been left alone to make the conceptual connections between the disciplines (National Research Council, 2012). Now, with these concepts as a central piece to the NGSS, teachers and students are provided the opportunity to give them due instructional time.

The concepts included in The Framework were selected because of their crosscutting nature. To qualify, concepts needed to be valuable organizing elements throughout the disciplines of science and engineering. Seven concepts were selected: Patterns; Cause and effect; Scale, proportion, and quantity; Systems and system models; Energy and matter; Structure and function; and Stability and change. The first two are powerful concepts in the nature of science. Scientific practice always includes pattern seeking and cause-and-effect relationships. The third concept, dealing with scale, makes the all-important connection between science, engineering, and mathematics. The final four concepts are related. The first, systems, is further explained through understanding the next three; however, all four are presented separately because the concepts of energy, structure and function, and stability and change are also important concepts in their own right (National Research Council, 2012).

These concepts should not be taught in isolation—that would defeat their very purpose. It is the integration of these concepts in the teaching of disciplinary core ideas that demonstrates their powerful value. As students have opportunity to explore content through the lens of these concepts across the disciplines, they provide a connective

structure that facilitates comprehension of what is being studied and how it relates to ideas in other disciplines (National Research Council, 2012).

Disciplinary Core Ideas

A challenging demand and key feature to The Framework is the need for cutting down the amount of content presented in the standards. Therefore, the content was built around disciplinary core ideas. Such packaging shifted the focus from teaching “all the facts” to building a foundation of core knowledge that would best prepare students to become critical consumers of science as well as continue in learning science, apply scientific understandings to life, and even become scientists (National Research Council, 2012). These core ideas were treated with fidelity in the development of the NGSS.

The Framework employed four criteria in establishing inclusion of a disciplinary core idea. Preferably the ideas that were accepted would meet all four criteria. However, this expectation was too lofty, and so the standard was set at two.

1. Have broad importance across multiple sciences or engineering disciplines or be a key organizing principle of a single discipline.
2. Provide a key tool for understanding or investigating more complex ideas and solving problems.
3. Relate to the interests and life experiences of students or be connected to societal or personal concerns that require scientific or technological knowledge.
4. Be teachable and learnable over multiple grades at increasing levels of depth and sophistication. That is, the idea can be made accessible to younger students but is broad enough to sustain continued investigation over years (National Research Council, 2012, p. 31).

The disciplinary core ideas were then grouped into major domain associations.

The first domain, physical sciences, combines traditional disciplines of physics and chemistry. Here the overarching goal is that students would see the mechanisms of cause

and effect which can be understood from an essential set of physical and chemical principles. The second domain, life sciences, extends to a study of all life on earth through unifying ideas. This is done with a focus on patterns, processes, and relationships. The third domain, earth and space sciences, is formed from a combination of many science disciplines and is strongly connected to both of the first two domains. Here geology is a cornerstone branching into astrophysics, geophysics, geochemistry, and geobiology. This domain investigates the processes of earth as well as its place in the universe. The fourth and final domain, engineering, technology, and applications of science, is included in The Framework as yet another attempt to ensure that the NGSS will not simply become another set of facts for students to learn. The goal of the core ideas in this domain is to progress students' understanding of the development of scientific knowledge and explanations as well as the practical applications of that knowledge (National Research Council, 2012).

Integration and The Framework

Integration is an underlying theme that runs through The Framework. Building from the integrated nature of science—including both intradisciplinary and interdisciplinary elements—multiple aspects of the domain are woven together in the framework. Each of these will be explored in this section of the review. However, integration is a complex topic and so the review will begin by delineating the relevant aspects of integration. Finally, these aspects of integration and the nature of science will be fit to The Framework and used to demonstrate connections between The Framework and the CCSS.

Relevant Aspects of Integration

Integration has been a topic of discussion within education since at least the turn of the 20th century (Mathison & Freeman, 1997), yet it has proven to be a difficult domain to define (Applebee, Adler, & Flihan, 2007; Shoemaker, 1991; Shriner, Schlee, & Libler, 2010). In this review three relevant aspects of the topic are explored and operationalized. Predominately, integration in education is associated with the combining of subject areas. This aspect will be reviewed first. Second, integration is viewed as the blending of school based knowledge and real-world application. This aspect is also relevant to the discussion of science education. Finally, from a perspective of cognitive science, integration is the assimilation of new information into the structures of the brain. Each of these three aspects of integration will be addressed in turn.

Integration has been a difficult term to define in the world of education, yet broad attempts have been made with some uniformity. Gehrke (1998) applied her broad definition to the domain, “it is a collective term for those forms of curriculum in which student learning activities are built, less with concern for delineating disciplinary boundaries around kind of learning, and more with the notion of helping students recognize or create their own learning” (p. 248). Case’s (1991) theoretical work contains similar elements and practicing educators agree with these broad statements about integration (Nollmeyer, 2013). This congruence centers on the idea that integration is about combining different subjects in the same lesson or unit. Therefore, this aspect of integration needs to be explored.

A wide variety of suggested models can be found for integrating subject areas. Fogarty (1991) describes ten ways that integration can occur; most of these are subject area based. She presents these along a linear continuum from less to more integration. The first, and least integrated design, she calls the *Fragmented Model* where each subject becomes the focus in turn. It looks much like traditional curriculum and the connections between the subjects are only implied. Towards the far end of Fogarty's continuum is the *Integrated Model*. Here, all four major school subjects blur as concepts, skills, and attitudes from each subject are brought together in a kaleidoscope approach. The integration occurs by following the related ideas that arise from within the disciplines. Viewing options in subject area integration as a continuum finds much agreement within the literature (i.e. Huntley, 1998; Leung, 2006; Lonning, DeFranco, & Weinland, 1998).

Integration of subjects includes as aspect of depth as well. The literacy across the curriculum movement at the end of the 20th century emphasized the use of literacy skills in the content area as a way to improve learning (Harp, 1989; McKenna & Robinson, 1990). Such integration contains a low level of depth (Nollmeyer, 2013) and would sit toward one extreme on Huntley's (1998) unique dual-subject continuum. Increased depth of integration is displayed in moving towards the middle of her continuum where each subject plays a more essential role in the learning. She says that *Integrated* curriculum (the center of her continuum) is where the teacher explicitly integrates concepts and skills from multiple subjects and gives equal time to both. Teachers in Nollmeyer's (2013) study described this depth of integration as meeting standards and having lesson

objectives from multiple subjects instead of simply using one or more subject areas as tools in enhancing the learning of a central subject.

Building subject area integration around science education is a common strategy in curricular integration at the elementary level. Greenleaf et al. (2011) promote this strategy for its dual benefits of strengthening literacy skills and acquiring deep understandings in science. Kovalik's model for designing curricular integration is a conceptually built integration that begins with science content (Greene, 1991). Even teachers with little or no explicit training in curricular integration prefer to design their units this way (Nollmeyer, 2013). Over the past two decades, Romance and Vitale (2001, 2011; 2006) have demonstrated that such a curricular design is successful. They conducted longitudinal studies that tracked student success in both English language arts and science standardized testing as well as persistence of learning.

Another aspect of integration is situating it in real-world situations and applications. The value of integrating school based knowledge and the real world is not a new idea. Dewey and other early progressivists promoted a real-world, experience oriented curriculum where students could connect school and life because their educational experiences would provide the connections explicitly (Bunting, 1987; Dewey, 1902/1966, 1938/1988; Tanner, 1989). Whitehead (1929) saw a life-based curriculum as essential in producing a meaningful education, "The solution which I am urging is to eradicate the fatal disconnection of subjects which kills the vitality of our modern curriculum. There is only one subject matter for education, and that is Life in all its manifestations" (p. 6). This was how Hollis Casewell wrote curriculum at his

laboratory in the 1930s. He worked with teachers and state committees to create curriculum based on themes from life like “conservation” or “consumption” (Fraley, 1977). The Progressive Education Association’s project that resulted in the Eight-Year Study was also framed by this ideal (Mitchell, 1985; Vars, 1991).

An integrated curriculum that is applicable to real-world situations is still a key characteristic of the domain today. This is not surprising considering the purpose of integration. When breaking down the artificial walls created by disciplines, the result is an organic situation where knowledge and skill are accessed as needed and synthesized to meet problems (Ledoux & McHenry, 2004; Mathison & Freeman, 1997). Connecting the content learned to student prior knowledge and current experiences is essential in building student motivation (MacMath, Roberts, Wallace, & Chi, 2010).

Three aspects of integration are relevant to this study; having addressed subject area integration and real-world integration, this review now turns to elements of integration in learning theory. Though cognitive science is relatively new, insights into the integration of knowledge within the mind can be traced to the mid-1800s and the writings of Herbert Spencer. The British psychologist suggested that the last step of a changing or adapting organism was that of integration. Whatever had changed must be successfully integrated into the whole organism for the adaptation, the evolution, to be a success (Mathison & Freeman, 1997). Viewing the term integration in this way is taking it at face value. It is the unifying of discrete elements into a whole (Case, 1991). Fifty years later Gestalt psychology connected Spencer’s thoughts on natural science to the realm of psychology.

Cognitive psychology traces its origins to the study of Gestalt psychology in the early 20th century. The theory of Gestalt psychology proposes that an internal integration occurs as a whole is interpreted. Piaget and Inhelder (1969) described learning in terms of assimilation or reconstruction. This restructuring creates new frameworks for understanding information in an updated fashion (Rumelhart & Norman, 1981). These descriptions of the interaction of prior knowledge and new information amount to a restating of Spencer's claim that integration is the last step in adaptation. This is just how Case (1991) describes the event, "The very act of learning typically involves integration—new beliefs are filtered through and connected to the individual's prior beliefs" (p. 215). However, facilitating this integration of knowledge requires bringing active student experience and large concepts together in unified moments (Beijaard & Verloop, 1996).

The integration of knowledge is also evident in the descriptions of specific cognitive theories. Information Processing theorists explain that these brain structures that have been formed become essential to the processing and assimilating of new information (Schunk, 2012), and Cognitive Load researchers demonstrate the power of prior knowledge to mediate the problem of load in the working memory (Paas, Renkl, & Sweller, 2003; van Merriënboer & Sweller, 2005). Joan Fulton's Cognitive Process of Instruction (CPOI) follows a seven step process through which teachers assist students in developing a "coding match" or conceptual structure between prior knowledge and the topic of study. This conceptual framework then becomes the students' model for thinking about situations that align with it (Fulton, 1989). Ledoux and McHenry (2004)

used CPOI with their preservice teachers while having them write integrated curriculum. They state that, “If learners are continuously attempting to make sense of environmental stimuli that are not divided according to false disciplinary lines, then the development of coding systems should match the authentic nature of the world, a world where there are no boundaries separating the processes of language arts from the content of science” (p. 391).

The integration of knowledge in the mind is discussed by Fogarty (1991) as she describes her *Immersed Model* of integration. The integration in this model, the ninth of her ten, is an internal integration occurring within the student. “In an intensely personal way, it filters all content through the lens of interest and expertise” (p. 64). And while Fogarty uses examples like a student completely immersed in one topic or a scientist completely focused on his research, Beane (1992) extends this to all students and all learning.

Authentic and significant learning occurs as new experiences are integrated into our scheme of meanings in such a way that those meanings are expanded and extended. The “integration” is done by the person him or herself; it is not done for that person by others. Moreover, this ongoing process begins with meanings already held, rather than from some abstract point that may eventually connect with them. (p. 49)

His point aligns well with Dewey’s (1902/1966) thoughts, “Subject-matter never can be got into the child from without. Learning is active. It involves reading out of the mind. It involves organic assimilation starting from within” (p. 9). This active learning is a process of assimilation and reconstruction, a combining of prior knowledge and new information, a creation of a new coding system for interpreting the world. It is integration.

The Intradisciplinary and Interdisciplinary Nature of Science

Science is a practice involving the effort of multiple disciplines. This includes Intradisciplinary interactions as well as interdisciplinary connections. However, before addressing these aspects, this review will operationalize these terms. Intradisciplinary refers to a consideration of a single discipline. For example, an intradisciplinary curriculum is one that teaches a single discipline (Huntley, 1998). Here disciplines are viewed as a self-contained domain of human experience with shared goals, concepts, skills, and judgments (Nissani, 1995). While, at the purest level this may separate the various fields of science, this research proposes that in light of a school curriculum, it is better to view the relationships between the fields of science as intradisciplinary interactions. Interdisciplinarity is connecting elements of two or more distinct disciplines (Nissani, 1995). For example, an interdisciplinary curriculum would process concepts and elements of multiple disciplines at one time. It may use common themes or processes to form connections (Mathison & Freeman, 1997).

The intradisciplinary connections in science can be discussed through the lens of the debate over domain specific versus domain general scientific thinking is one demonstration of the intradisciplinary nature of science. Domain-specific approaches to scientific thought emphasize the need for conceptual understanding within a domain in solving problems. Domain-general approaches emphasize the development of problem-solving strategies that can be used across the disciplines of science (National Research Council, 2007; C. Zimmerman, 2000). Klahr and Dunbar (1988) presented an integrated view of these approaches as a comprehensive approach to the discovery process which

they termed Scientific Discovery as Dual Search (SDDS). Through a number of studies they demonstrate the value of understanding scientific discovery from this dual, integrated perspective. The model demonstrates the need for both domain-general strategies and domain-specific knowledge in scientific practice (C. Zimmerman, 2000).

Science is interdisciplinary because of its influence in human life and activity and its necessary connections to other school subjects. “Because scientific knowledge is produced in a social context it is necessarily impacted by the personal and professional goals of scientists, the interests and priorities of funding agencies, and the cluster of economic, political and moral-ethical influence that impregnate the sociocultural context in which scientific practice is located” (Hodson, 2009, pp. 151-152). Every aspect of life is influenced by scientific discovery; therefore, stakeholders come from every disciplinary domain (National Research Council, 2012).

The nature of science also includes interdisciplinary connections to mathematics and language arts. Scientists regularly employ mathematics in organizing and analyzing data as well as in formulating explanations, designing models, and being critical consumers. They also must be able to both read the work of others and communicate effectively about their own work (National Research Council, 2012).

The Framework Features that Promote Integration

Each of the discussed aspects of integration and the interdisciplinary nature of science are woven into The Framework. The Framework involves subject-area integration by linking the various disciplines of science, including the knowledge and practice of engineering, and making connections to the CCSS. The crosscutting concepts

that form one leg of the three-legged framework for The Framework have built intradisciplinary integration into the fabric of the standards. Yet, the NRC (2012) expects more. They say that continuity across the subject areas is vital to student success since these connections provide opportunities for reinforcement and application. These connections need to remain sensible and not be contrived (National Research Council, 2007). The interdisciplinary integration of engineering into The Framework was a natural one since both permeate every aspect of life. And while their practices differ slightly, both require a combination of knowledge and skill and both can be viewed through the “Three Spheres of Activity” model (see Figure 1) (National Research Council, 2012). Finally, a special effort to integrate the ELA and mathematics CCSS into The Framework has been made (National Research Council, 2012; NGSS Lead States, 2013).

The Framework also promotes real-world integration through its emphasis on scientific practice and relevant application of scientific knowledge. Instead of isolating “skills” from content and teaching them as a single set of step-by-step procedures (i.e. The Scientific Method), The Framework expects students to participate in genuine scientific practice which always situates knowledge and skill in the natural world—exploring, testing, formulating explanations and models. The Framework also expects students to solve (albeit not with the knowledge base of an adult) real-world problems by applying engineering knowledge and skill (National Research Council, 2012; NGSS Lead States, 2013). For example, “3-PS2-4. Define a simple design problem that can be solved by applying scientific ideas about magnets. [Clarification Statement: Examples of

problems could include constructing a latch to keep a door shut and creating a device to keep two moving objects from touching each other.]” (NGSS Lead States, 2013, p. 25). This performance expectation not only gives students an opportunity to solve a real-world problem, it also allows them to see how scientific discovery and understanding provides engineers with the necessary information (NGSS Lead States, 2013).

Finally, one of the three pillars of The Framework is the Crosscutting Concepts. These concepts bridge the disciplinary gap between the sciences. Their inclusion mirrors the understanding of integration in cognitive science. The cross-cutting concepts are to act as a conceptual framework that allows students to develop new understandings and make linkages across the disciplines. These concepts can succeed in such a task because of their innate pervasiveness (National Research Council, 2012). Current understandings in science, regardless of their naivety, are crucial in building new understandings (National Research Council, 2007). The NRC (National Research Council, 2012) have taken this immensely personal process (Fogarty, 1991) and provided educators with concepts to assist students in the proper development of a scientific worldview (National Research Council, 2012).

Developing The Framework has been a comprehensive effort and produced a complex product. Determining the readiness of practicing elementary teachers to implement these standards will require an instrument that operationalizes this implementation and measures the various factors with validity and reliability.

Instrument Design

This portion of the review explores the literature base for instrument design and scale development. Successful instrument design begins with a clear understanding of what is being measured. It is built with validity and reliability in mind, and demonstrates those qualities under statistical analyses. The next sections consider each of these factors in turn.

Describing what is Being Measured

Clarity about what is being measured requires clarity about the latent constructs within the phenomena. It is not uncommon for researchers to think that they are clear on the purpose of their instrument only to discover that they have failed at that very point (DeVellis, 2003). Researchers must develop a precise and comprehensive understanding of the constructs being measured (Clark & Watson, 1995) which can be thought of as the test framework (American Educational Research Association, 1999). The precision in description includes defining each construct with an appropriate amount of specificity. Underrepresentation of a construct can well be the beginning of invalidity (Netemeyer, Bearden, & Sharma, 2003). Factors resulting from exploratory analysis should only be considered further if at least three items load exclusively on the factor (MacCallum, Widaman, Zhang, & Hong, 1999); therefore, fully operationalizing each construct is essential and is why consulting the literature base for conceptual frameworks and definitions is so important in instrument design (Fink, 2003).

A construct is a characteristic that is not always directly observable but measureable through scale items. It can be a personal attribute (Cronbach & Meehl,

1955) or a conceptual phenomenon of interest (Edwards & Bagozzi, 2000). Edwards and Bagozzi (2000) explain that while researchers “construct” these concepts they are none-the-less real; therefore, constructs can be assessed through measurement (Clark & Watson, 1995; Cronbach & Meehl, 1955; DeVellis, 2003; Edwards & Bagozzi, 2000; Netemeyer et al., 2003). In fact, scale items are often the best way to assess present constructs (DeVellis, 2003) because theory on constructs and measures state that variation in constructs will lead to a variation in measures (Edwards & Bagozzi, 2000). However, it is essential that the constructs being measured are grounded in theoretical literature. Without such a foundation, little relevance can be assured (Netemeyer et al., 2003).

DeVellis (2003) explains that the assumption by the researcher is that measures are a “proxy” for constructs. A relationship between scale items demonstrates a similar relationship between the latent constructs with which each item is associated. Depending upon the presence of such a relationship further emphasizes the need for theoretical grounding (Netemeyer et al., 2003). As may be assumed, more helpful scale results tend to be produced by efforts involving a deeper theoretical understanding (Clark & Watson, 1995).

Ensuring Validity

The validity of a survey is the notion that the instrument measures what it says it measures. An instrument is valid if the latent variables are indeed the cause for the observed variance (DeVellis, 2003; Netemeyer et al., 2003). Since validity is the most fundamental issue in test development (American Educational Research Association,

1999), scale developers need to carefully address the different types of validity: Content, Criterion, Construct, and Consequential.

Content validity refers to the extent to which instrument items adequately measure the characteristics of a domain (American Educational Research Association, 1999; DeVellis, 2003; Fink, 2003; Haynes, Richard, & Kubany, 1995) or how well a particular set of items measure an individual construct (Lux, 2010). It is not enough that items accurately address aspects of a construct. Content validity hinges on how well items address all characteristics of a construct (Netemeyer et al., 2003). The “adequacy” of a particular argument for validity changes depending upon the application of the results. For example, adequate content validity for a screening instrument may be woefully inadequate for valid determination of a treatment plan (Haynes et al., 1995). In order to maintain content validity, researchers write items for each aspect of their carefully operationalized constructs—which in turn are often built upon theoretical literature (Fink, 2003). At times, this validation has a reverse affect where constructs go through definitional adjustments; here content validity blurs with construct validity (American Educational Research Association, 1999; Haynes et al., 1995).

Criterion validity refers to the extent to which items have an empirical relationship with a criterion (DeVellis, 2003). As such, it is a predictive validity and often so named (Fink, 2003; Netemeyer et al., 2003). It is not concerned with theoretical relationships or causal relationships and emphasizes more of a practical than scientific reality (DeVellis, 2003). If a meaningful correlation is observed then criterion validity

has been established. The stronger the correlation observed the stronger the validity (DeVellis, 2003; Netemeyer et al., 2003).

Construct validity refers to the extent to which an instrument successfully distinguishes between participants during experimentation (Fink, 2003). This distinction occurs if the theoretical relationship between a participant's beliefs and the instrument items is born out (DeVellis, 2003). With increased demands on the validity of published instruments, construct validity has become vitally important (Clark & Watson, 1995) and is especially needed when there is no established, agreed upon definition of a construct (Cronbach & Meehl, 1955). Construct validity can be confused with criterion validity since correlations are used in establishing both; however, the purpose of the correlation makes all the difference (DeVellis, 2003). Plus, construct validity is not based upon one set of observations, but rather, a series of tests (Clark & Watson, 1995).

Consequential validity is the extent to which an instrument serves its intended purpose (Birenbaum, 2007). Bangert (2009) explained that a diagnostic reading instrument would have consequential validity if it resulted in accurate placement of students in remedial reading programs.

Ensuring validity is not simply a matter of demonstrating the validity of one part of the test. An integrated argument must be made from all of the available evidence including information about test construction, reliability scores, procedures, score scaling, standard setting, and fair treatment of participants (American Educational Research Association, 1999).

Ensuring Reliability

Reliability refers to the consistent performance of a scale instrument. When the test is repeatedly administered with individuals, how consistently does it score the same participant (American Educational Research Association, 1999)? Therefore it is related to the true-score; however, since the true-score cannot be observed, reliability must be inferred from the observed score through the use of one or more of three types of reliability: test-retest, alternative-form, and internal consistency (Netemeyer et al., 2003). Test-retest reliability refers to the strength of correlation from one administration to the next (Fink, 2003; Netemeyer et al., 2003). Alternative-form reliability, also called equivalence, refers to the extent to which the level of difficulty remains uniform across different forms measuring the same constructs (Fink, 2003). This uniformity is established through correlation (DeVellis, 2003; Netemeyer et al., 2003). Internal consistency refers to the extent to which items measuring a construct are correlated with each other (DeVellis, 2003; Fink, 2003; Netemeyer et al., 2003). It is this last form of reliability that is used most frequently (Netemeyer et al., 2003).

While the relationship between scale items and their latent construct cannot be directly observed, correlation between scale items is an indicator for each item's correlation to the same latent variable (DeVellis, 2003). The most common measure used to evaluate these correlations is coefficient alpha (Netemeyer et al., 2003). In fact Cortina (1993) argues that the development of alpha was one of the most important developments in instrument design. In designing and sampling an instrument, the

researcher has attempted to prune weak items and keep strong ones. The coefficient alpha is the most important measure of success in this endeavor (DeVellis, 2003).

In presenting coefficient alpha, Cronbach (1951) described the need met by his alpha. First, the best way to establish internal consistency is to compare two independent measures; however, there are nearly always practical challenges to this. Second, the traditional split-half has faced multiple criticisms (i.e. Brownell, 1933; Kuder & Richardson, 1937) (as cited in Cronbach, 1951). From this backdrop Cronbach (1951) then posed this question, “The essential problem set in this paper is: How shall α be interpreted?” (p. 300). His answer was coefficient alpha or as it is commonly called, Cronbach’s alpha (Cortina, 1993; Netemeyer et al., 2003).

Variation in a set of item scores can be attributed to two things. Either it is actual variation of participants in relation to the construct or it is error (DeVellis, 2003). Coefficient alpha takes the total variance and separates it into these two categories so that $1 - \text{error variance} = \alpha$ (DeVellis, 2003; Netemeyer et al., 2003). Cortina (1993) demonstrates the value of coefficient alpha over other measures of internal consistency by pointing out that since Cronbach’s alpha is “lower bound,” unlike standardized alpha, it can be further trusted to represent true linearity in variance because,

As the items in tests approach essential tau-equivalences (i.e. linearly related and differing only by a constant), as they do when the tests are composed of equal portions of general and group factor variance, Cronbach’s alpha approaches reliability. When test items are exactly essentially tau-equivalent, Cronbach’s alpha equals reliability. (p. 101)

Regardless of the alpha calculation used, it is important to remember that they all are influenced by the number of items in a scale (Cortina, 1993; DeVellis, 2003).

Because of this relationship, Cortina (1993) warns against judging scores according to the

most common standard of $\alpha = .70$. Most studies use this number regardless of scale size. While DeVellis (2003) states that he is most comfortable with $\alpha = .70$ to $\alpha = .80$ as a respectable range and $\alpha = .80$ to $\alpha = .90$ as a very good range, he also addresses the influence of the number of items on the alpha score. “A scale’s alpha is influenced by two characteristics: the extent of the covariation among the items and the number of items in the scale” (DeVellis, 2003, pp. 96-97). The problem then is that far too often, readers assume that seeing a score like $\alpha = .70$ can automatically be interpreted as adequate without any further discussion (Cortina, 1993). Cortina presents this example of the danger in such reading.

As an example, I compare the meaning of standardized $\alpha = .80$ for scales made up of 3 and 10 items. For a 3-item scale with a $\alpha = .80$, the average interitem correlation is .57. For a 10-item scale with $\alpha = .80$, the average interitem correlation is only .28. This is strikingly different from .57 and underscores the fact that, even without taking dimensionality into account, alpha must be interpreted with some caution. (p. 101)

This is not to say that longer is always better (even if you can get participants to complete very lengthy surveys). Scales that are too long bloat alpha scores artificially (Cortina, 1993; DeVellis, 2003). For this reason, DeVellis (2003) suggests a shorting of the scale anytime the score is much over $\alpha = .90$.

Themes for The Framework

Developing an instrument to measure the *understanding* and *readiness* of elementary educators to implement The Framework, required the identification of the themes associated with a successful implementation. Themes characterizing The Framework were identified through a review of relevant literature. In particular, the

NRC's (2007) *Taking Science to School: Learning and Teaching Science in Grades K-8*, The NRC's (2012) *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*, and the resulting *Next Generation Science Standards: For States, By States* (NGSS Lead States, 2013).

Six themes were identified through the process of the review: (1) Scientific and Engineering Practices; (2) Crosscutting Concepts; (3) Disciplinary Core Ideas; (4) Integration of Three Dimensions; (5) Best Practices in Science Instruction; and (6) Connections to Common Core. Edwards and Bagozzi (2000) propose that constructs must be conceptually defined and Netemeyer et al. (2003) state that the importance of such definitions cannot be over emphasized. And, "We have found that writing out a brief, formal description of the construct is very useful in crystallizing one's conceptual model" (Clark & Watson, 1995, p. 310). To that end, formal descriptions of each are presented here.

Scientific & Engineering Practices

In the past, this section might have been termed, "Science Inquiry" or "Science Skills." The NRC (2012) intentionally avoided these terms to emphasize the integration of skill and knowledge that occurs in scientific and engineering practices. Scientific inquiry is a crucial scientific practice, but the term "practice" is more inclusive of all the activities in which scientists and engineers engage (Michaels et al., 2008). Instead of merely an additional aspect of science education (or a set of steps to follow—i.e. The Scientific Method), these practices replicate how science works. It is not surprising then that this aspect of science education includes elements of all four of the NRC's (2007)

strands of scientific proficiency: know and interpret explanations, generate and evaluate evidence, understand the nature of knowledge, and participate in practice.

Eight scientific and engineering practices are essential elements in science education. These are activities that scientists and engineers engage in as part of their jobs.

1. Asking Questions (for science) and defining problems (for engineering)
2. Developing and using models
3. Planning and carrying out investigations
4. Analyzing and interpreting data
5. Using mathematics and computational thinking
6. Constructing explanations (for science) and designing solutions (for engineering)
7. Engaging in argument from evidence
8. Obtaining, evaluating, and communicating information. (National Research Council, 2012, p. 49)

Crosscutting Concepts

Crosscutting concepts act as a bridge between the disciplines of science and engineering. They have value in explaining phenomena across the disciplines and therefore act as an framework for students to organize their understandings and enhance their scientific view of the world (National Research Council, 2012). These are not contrived linkages, they naturally infiltrate the various domains and have done so throughout history. These concepts transcend disciplinary bounds and therefore provide great value in explanation, theory, observation, and design (American Association for the Advancement of Science, 1990). Since educators cannot trust children's naïve views on science to be naturally replaced with the passing of time, educators must be more intentional. This includes sequencing of new information and connections between

concepts being taught (Duschl, 1990). These connections will not be built on their own—as history has shown (National Research Council, 2012).

Seven crosscutting concepts have been selected based upon their far-reaching value throughout the disciplines of science and engineering: (1) Patterns, (2) Cause and effect, (3) Scale, proportion, and quality, (4) Systems and system models, (5) Energy and matter: flows, cycles, and conservation, (6) Structure and function, (7) Stability and change (National Research Council, 2012; NGSS Lead States, 2013).

Disciplinary Core Ideas

The need to trim the volume of content in national science standards is well documented (i.e. Bybee, 2006; Coleman & Zimba, 2008; National Research Council, 2007; Sneider & Workosky, 2009). Therefore, the NRC (2012) greatly reduced the number of details to be learned at a shallow level and instead opted for a deeper, more thorough investigation of “Core Ideas.” The result of this conceptual versus factual organization was that four core ideas were identified for the physical sciences, four for the life sciences, three for the earth and space sciences, and two for engineering, technology, and applications of science (National Research Council, 2012; NGSS Lead States, 2013).

Integration of Three Dimensions

The three dimensions of The Framework do not stand alone: Scientific & Engineering Practices, Crosscutting Concepts, and Disciplinary Core Ideas. Instead, these three distinct aspects must also be inseparable. “[Students should] actively engage in scientific and engineering practices and apply crosscutting concepts to deepen their

understanding of the core ideas in these fields” (National Research Council, 2012, pp. 8-9). This integration is first seen in the NRC’s (2012) choice to avoid the term “skills” in favor of “practices.” Traditionally, skills may have been seen as separate and distinct. When scientists or engineers practice they are combining conceptual frameworks of understanding along with skills to inquire about the natural world. This then is a model for how knowledge and skill should be intertwined in the classroom as well. Teaching “science as practice” requires integration (Michaels et al., 2008). The obvious integration of the crosscutting concepts is in the very nature of their presence—they cut across the domains of science and engineering. Evidence of this should be seen in their repeated use in the processes of instruction and practice. They should act as a framework connecting knowledge with knowledge and knowledge with practices. This will be a challenge practically; “A major question confronting each curriculum developer will be which of the practices and crosscutting concepts to feature in lessons or units around a particular disciplinary core idea, so that, across the curriculum they all receive sufficient attention” (National Research Council, 2012, p. 247). Yet, this challenge has been mediated by the NGSS document which presents an integrated picture of the three dimensions in each Performance Expectation (NGSS Lead States, 2013).

Best Practices in Science Education

When students learn they do so by construction. They assimilate new knowledge and skill into prior understandings or reconstruct current structures, integrating new knowledge and skill into an updated framework of understanding (Piaget & Inhelder, 1969). This necessitates an active participation on the part of the students (Collins,

2002). The Framework expects students to be active in their learning. The integration of the three dimensions and the linkages provided by the crosscutting concepts are both built upon the idea that the brain is actively integrating knowledge (National Research Council, 2012).

How this actually looks is quite complex. Previous debates about direct instruction versus inquiry oversimplify the issue. Science instruction should involve a wide range of scientific activities and thinking including inquiry, gathering and evaluating evidence, logic, communication, and application of knowledge (National Research Council, 2007, 2012). Some of these will be teacher-led activities, and others will be student-led activities. There will be direct instruction, teacher questioning, and collaborative small-group investigation—all monitored carefully through assessments and adjusted according to the need for scaffolding (National Research Council, 2012).

Connections to Common Core

There are multiple connections between the new CCSS (both mathematics and ELA) and The Framework. The title of the Common Core Standards for English Language Arts is *English Language Arts and Literacy in History/Social Studies, Science, and Technical Subjects*. This title makes it clear that ELA skills are a necessary element in understanding. The new standards demand a high level of reading competency and bring back an emphasis on writing (Gewertz, 2012) which matches a fundamental quality of science practice. Both written and spoken communication is essential in scientific practice. Scientists must describe with precision, present explanations with clarity, and justify through argumentation. They must also be critical readers of other's work. These

are practices that students must develop as they engage in the activities of scientists in the field (National Research Council, 2012).

The Common Core mathematics standards also have strong connections to The Framework. Scientists need to have the ability to handle and interpret quantitative data. They organize, tabulate, graph, and statistically analyze data to uncover meaning. Students need the opportunities to participate in these practices both in the creation and organization of data as well as in the exploration of available data (National Research Council, 2012). Connecting The Framework with Common Core has been made easier for the practicing educator by the resulting NGSS document which provides connections for both disciplines (NGSS Lead States, 2013).

The Constructs of Understanding and Readiness

Understanding and readiness have connections with each other, yet they are unique and unequal constructs. In his reflection on the revision of Bloom's Taxonomy, Krathwohl (2002) explains that the comprehension category in the original taxonomy was renamed "understand" in the updated list. However, he adds that the reason for its absence in the original taxonomy was that it is too frequently used to refer to the strength of student learning regardless of the Bloom's category. This is the way in which the word is used throughout the literature—connected to both general knowledge and deep knowledge. It is connected with big conceptual ideas over the knowledge of discrete facts (i.e. Grant, 1999; Sackes, 2010; Wavering, Mangione, & McBride, 2013). Understanding is also used to refer to complete, essential, or full knowledge of a topic (i.e. Cisterna, Williams, & Merritt, 2013; National Aeronautics Space Administration,

1998). In short it is the go to word, for learning. So, “When teachers say they want a student to ‘really’ understand, they mean anything from *Comprehension* to *Synthesis*” (Krathwohl, 2002, p. 214).

Readiness to implement ideas or readiness to make changes is connected to what is understood, yet fully describing all the variables involved is a complex task (Steele, Brew, Rees, & Ibrahim-Khan, 2013). Readiness goes beyond understanding because it contains the expectation of application. At its heart, readiness is not about cognitive knowledge only (as understanding can be described) but rather practical knowledge (Beijaard & Verloop, 1996). Beijaard further explains, though, that this does not mean that readiness is devoid of understanding. On the contrary, readiness is the application of understandings to a relevant situation via adaptation. At the same time, it is also influenced by emotion and self-efficacy. Research has found that positive feelings about technology were strongly correlated with technological readiness (Parasuraman, 2000). In science education, Steele et al. (2013) found that readiness to teach science was most strongly influenced by prior experiences with science as a student. Positive experiences with science in the past produced higher levels of confidence and enthusiasm about teaching science than any other variable. This self-confidence and belief in ability is called self-efficacy (Bandura, 1986). Low self-efficacy results in limited and poor science instruction (Scharmann & Orth Hampton, 1995).

Chapter Summary

The history of science education situates current reforms in light of historical efforts, ideas, attempts, successes, and failures. This review presented a current and

unsatisfactory status of science education. Then, it traced the history of science education—highlighting the progress of and often disparate realities between theory and practice. Throughout this discussion, the influences of the nature of science, philosophical and psychological ideas, and policy reforms were considered. Finally, the review returned to the current state of science education with a view to the future. The works of the NRC (1999, 2007) were considered at some length as they form a foundation for The Framework. A review of the literature on instrument design and the operationalizing of the proposed themes closed the chapter.

CHAPTER THREE

RESEARCH METHODOLOGY

Introduction

The third chapter is devoted to describing the methods used during the course of this research. A number of topics are discussed to provide a full picture of the study from start to finish. The purpose of this study was to identify key constructs in *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Research Council, 2012) that could be used to develop an instrument, the *New Framework for Science Education Survey of Teacher Understanding & Readiness (NFSE-STUR)*, measuring elementary teachers' *understanding* of and *readiness* to implement The Framework.

The development of an instrument began with a thorough review of the literature related to elementary science curriculum and The Framework. The themes found in the literature were identified as, (1) Scientific and Engineering Practices; (2) Crosscutting Concepts; (3) Disciplinary Core Ideas; (4) Integration of Three Dimensions; (5) Best Practices in Science Instruction; and (6) Connections to Common Core.

These themes were used to guide the development of draft items which operationalize the themes for assessing elementary teachers' *understanding* and *readiness* for using The Framework to guide their instruction. The draft items were then distributed to experts for their review. Based on feedback from these experts, the survey was revised and pilot tested with a small group of teachers. Then, a final version of the survey was

constructed and administered to a regionally representative sample of elementary teachers. Finally, the data from the survey was analyzed for internal consistency and both an exploratory analysis and confirmatory factor analysis were conducted on each construct (*understanding* and *readiness*) in efforts to establish evidence for instrument validity. The steps followed during the development and validation of the *NFSE-STUR* were guided by recommendations from experts specializing in the field of instrument development (American Educational Research Association, 1999; DeVellis, 2003; Netemeyer et al., 2003).

Participants

There were three different sets of participants involved in this study. First, five individuals were identified as expert review panelists. These individuals were knowledgeable about current understandings of science learning as presented in the NRC's (2007) report, *Taking Science to School: Learning and Teaching Science in Grades K-8*, of the NRC's (2012) *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*, and the NGSS Lead State's (2013) *Next Generation Science Standards: For States, By States*. A second set of participants was involved in piloting the instrument. The goal was to have $N=20$ involved in this stage. These participants were students enrolled in an online graduate level course about the new framework for science education. They were all practicing elementary, middle, or secondary science teachers. The last set of participants in the study were involved in the final validation of the instrument. The goal was to have $N=300$ in this portion of the

study. They were all practicing elementary, K-5, teachers from the States of Montana, Idaho, Wyoming, and Utah.

Instrument Development

Creating the Item Pool

While the initial item pool was created based upon the review of the literature in totality, several documents were essential in developing the theme structure for organizing and writing items: the National Research Council's (2012) *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*; the NGSS Lead State's (2013) *Next Generation Science Standards: For States, By States*; and the National Research Council's (2007) *Taking Science to School: Learning and Teaching Science in Grades K-8*.

A structure of six themes emerged from the literature review to explain the aspects of science education as expected by The Framework. These six themes were scientific and engineering practices, crosscutting concepts, disciplinary core ideas, integration of the three dimensions, best practices in science education, and connections to the Common Core. Each item in the pool was written to assess teacher *understanding* and *readiness* in implementing one particular theme. A number of items were written for each theme. This redundancy is an important element in the creation of an item pool. Each item in the pool addressed either a decision or a behavior made by educators (Alreck & Settle, 2004) and began with this phrase: "When planning and teaching, educators" This phrasing was intended to emphasize the decisions and behaviors of educators who teach from The Framework with mastery. Instead of asking the

participants to compare their own practice to this standard, they were asked to rate their *understanding* of and their *readiness* to implement the idea. The prompts were written using two Likert scales to determine the magnitude of participant *understanding* and *readiness* for each item statement. During the pilot study the following four response options were provided for rating understanding: 1 = No Understanding, 2 = Slight Understanding, 3 = Fair Understanding, 4 = Strong Understanding. The following four response options were provided for rating readiness: 1 = No Readiness, 2 = Slight Readiness, 3 = Fair Readiness, 4 = Strong Readiness. These scales permit some flexibility in participant responses; while at the same time, not allowing for a “neutral” position to be taken on any item (Clark & Watson, 1995). The scales were lengthened to six-point scales for the validation study based on the need to more accurately identify variations. For the validation study, these options were provided for rating understanding: 1 = No Understanding, 2 = Slight Understanding, 3 = Fair Understanding, 4 = Solid Understanding, 5 = Strong Understanding, 6 = Advanced Understanding. These options were provided for rating readiness: 1 = No Readiness, 2 = Slight Readiness, 3 = Fair Readiness, 4 = Solid Readiness, 5 = Strong Readiness, 6 = Advanced Readiness.

To facilitate participant responses, clarifying quotes from *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Research Council, 2012) were connected to each prompt via a pop up window link labeled, “please click here for a detailed explanation.” Since the items were built from The Framework document and measured teacher *understanding* and *readiness* to implement The

Framework, it was appropriate to use these direct quotes to assist their self-reporting.

Figure 2 displays prompt 8 from the pilot survey. Figure 3 displays the pop up window quote associated with prompt 8.

8. When planning and teaching, educators have students organize, analyze, and interpret observational data to make meaning.

Please click here for a detailed explanation

Please rate your understanding of this idea.

No understanding	Slight understanding	Fair understanding	Strong understanding
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please rate your readiness to implement this idea.

No readiness	Slight readiness	Fair readiness	Strong readiness
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 2. Screenshot of Pilot Draft Item 8.

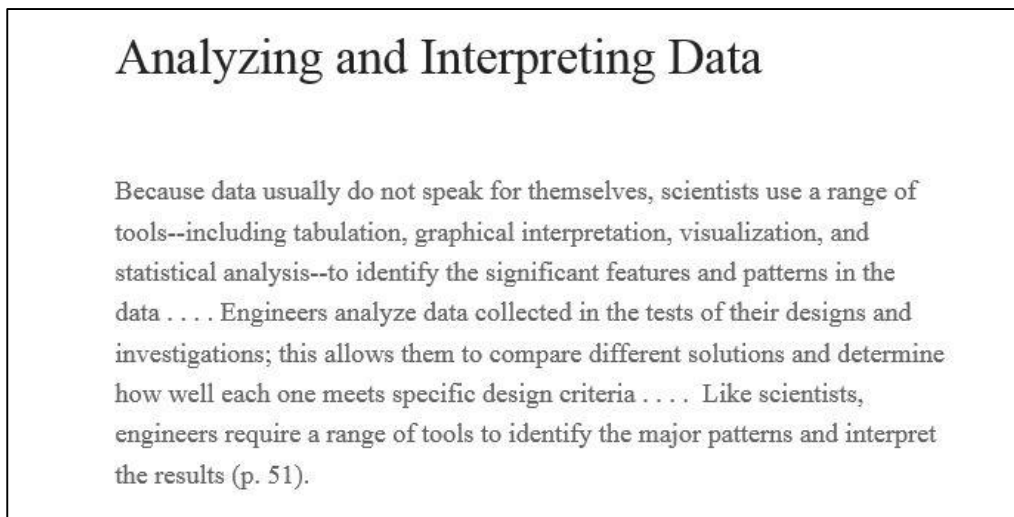


Figure 3. Screenshot of Framework Quote for Pilot Draft Item 8.

More items were created initially than would be needed or used in the final version of the validated instrument. This allowed the researcher the chance to select the best items from the pool (American Educational Research Association, 1999). Writing more items than needed allows for the creative exploration of the endless aspects of the

constructs and is recommended for the development of new instruments (DeVellis, 2003). Expert review and pilot testing provided the opportunity for removing extraneous items produced through this redundancy (Netemeyer et al., 2003).

Expert Review

A panel of experts was used to review the initial draft items (See APPENDIX A) developed by the researcher to assess the six themes identified from the literature. An expert review of a survey is an important step in instrument design as it assists in maximizing content validity (DeVellis, 2003). Experts, in the domain of study, are specifically suited to this task as they can recognize content quality and clarity (American Educational Research Association, 1999) and validate the framework for research and construct definitions (DeVellis, 2003).

Five experts in science education were approached about participating on the Expert Review Panel for the study. All five accepted the invitation and completed the review. Several factors were considered in identifying participant experts. Two key characteristics needed to be true for all participants. First, they needed to possess a high level of expertise in The Framework and NGSS as presented in the NRC's (2012) *A New Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Second, they needed to have a rich background in science education. With these two key characteristics as a base for selection, variables requiring diversity were considered. It was advantageous to have varying perspectives on elementary science instruction as well as varying perspectives on the Framework. To this end variety in the following characteristics were considered: regional or national experience, elementary

or secondary science instruction, professional development or higher education, and regional or national organization involvement. The five experts invited to participate brought rich diversity to the panel and provided a comprehensive perspective.

Dr. John Graves is a Faculty Lead Instructor in the Intercollege Programs for Science Education, Master of Science in Science Education and Montana State University, in Bozeman, Montana. John has over twenty years of experience as a middle school science teacher. He currently is the Executive Director of the Montana Science Teachers Association, recently served a four year term as the NSTA District XV Director. In the past, John has been the editor for the Montana Science Teachers Association News Journal and president of the Montana Science Teachers Association. He has significant professional development experience working as the Science Education Specialist for the Montana Partnership with Regions for Excellence in STEM grant and the Science Education Specialist for the Southwest Montana Math Science Partnership Project. John has a Bachelor of Science in elementary education from Eastern Montana College, a Master of Science in curriculum and instruction with a science emphasis, and a Doctor of Education in curriculum and instruction from Montana State University.

Paul Andersen has been teaching science in Montana for the last twenty years. Paul is currently a science teacher at Bozeman High School in Bozeman, Montana where he uses technology and guided inquiry to differentiate instruction for his students. Paul founded and maintains the website, www.bozemanscience.com which is dedicated to advancing best practices in science education. He has created hundreds of

science videos that have been viewed millions of times by teachers and students around the world. Paul was the Montana Teacher of the Year and a finalist for National Teacher of the Year. He has created a sixty video series detailing *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Research Council, 2012).

Ted Willard is a Program Director of NGSS@NSTA at the National Science Teachers Association (NSTA). In that role, he oversaw NSTA's feedback during the development of the *Next Generation Science Standards: For States, By States* (NGSS) and now coordinates NSTA's efforts to support teachers in implementation of NGSS.

Prior to joining NSTA two years ago, Ted spent 12 years at Project 2061. There he was responsible for the development of the growth-of-understanding maps published in the *Atlas of Science Literacy*, Volume 2. Ted also was involved in many other areas of Project 2061's efforts towards standards-based education reform including curriculum resources development, assessment development, science education research and teacher professional development.

Earlier in his career, Ted spent five years teaching high school physics in Asheville, North Carolina and five years editing elementary and high school science textbooks for the Globe Book Company (now part of Pearson) and Harcourt Brace School Publishers (now part of Houghton Mifflin Harcourt). Ted holds a Bachelor of Science in earth, atmospheric, and planetary science from the Massachusetts Institute of Technology.

Dr. Eric Brunsell is an Associate Professor of Science Education at the University of Wisconsin Oshkosh and a Core Instructor in the Intercollege Programs for Science Education, Master of Science in Science Education at Montana State University in Bozeman, Montana. Eric also has experience teaching as a high school science teacher and as director of education programs for a non-profit group focused on science teacher professional development. He has extensive experience facilitating professional development nationally and internationally. Eric is currently finishing his term as NSTA District XII Director. He is an active writer with articles appearing in NSTA journals, compilations from NSTA Press and a book from ASCD. His latest book project, focused on implementing professional development for the NGSS, should be released by NSTA Press early this summer. Eric has a Bachelor of Science in Secondary Science Education (Physics) from the University of Wisconsin Madison, a Master of Science in educational leadership with a technology integration emphasis from the University of Wisconsin Oshkosh, and a Doctor of Education in curriculum and instruction with an emphasis in science education from Montana State University.

Chris Campbell is a science teacher at Simsboro High School in Simsboro, Louisiana. He currently is the director for the NSTA District VII which encompasses the states of Arkansas, Louisiana, and Mississippi. He is working with the Lincoln Parish School District in coordinating the science curriculum for grades 3-5 as they prepare for the implementation of the NGSS in the 2014-2015 school year. He also works to train elementary teachers on incorporating science activities with literacy elements from CCSS. During the 2012-2013 year, Chris served as the Albert Einstein distinguished

educator fellow with the National Science Foundation. He holds a Bachelor of Arts in elementary education from the University of New Orleans and a Master of Science in curriculum and instruction from Louisiana Tech University.

The experts participated in the review through two stages. First, they were asked to review and rate items according to DeVellis' (2003) criteria: (1) Relevance of the items to constructs, (2) Items' clarity and conciseness, and (3) Point out missing aspects of constructs. Experts were also asked to rate and comment on the helpfulness of the framework quote. They did this through the online survey tool, surveygizmo.com (SurveyGizmo, 2014). In order to facilitate responses, the items were grouped on pages according to intended theme. Providing this assistance is important when expecting experts to comment on relevance (Lux, 2010). Figure 4 presents item 23 of the expert panel's first stage.

23. When planning and teaching, educators have students investigate phenomena in terms of structure and function as a means of sense making.

[Please click here for a detailed explanation](#)

Relevance: -- Please Select --

Clarity: -- Please Select --

Conciseness: -- Please Select --

NRC Quote Helpfulness: -- Please Select --

Comments: [Text Area]

Figure 4. Screenshot of Item 23 on the Expert Review Stage One Survey.

For each item, experts rated its relevance, clarity, conciseness, and the quote's helpfulness on a 4 point scale: 4 = Very Strong, 3 = Strong, 2 = Weak, 1 = Very Weak.

To ensure that the expert reviewers understood the task they were being asked to complete, an instructional video was created to accompany the written directions at the beginning of each page of the survey.

In preparing for stage two of the expert review, the researcher used the median scores for each rating to assist in evaluating the need for modification or removal of prompts; however, reviewer comments were found to be the most vital in determining the needs for change. Few changes, though, were actually made based solely upon data from stage one. Issues that arose during the experts' review of the prompts were compiled in a PowerPoint document (See APPENDIX B) and used during a focus group (consisting of the expert panel) webinar. Adobe Connect Software was used to conduct the webinar during which experts discussed comments from stage one and provided clarification for needed changes.

The expert review provided evidence for modifications, eliminations, and additions to the item pool. From the review the pilot draft of the instrument, *New Framework for Science Education: Survey of Teacher Understanding & Readiness (Pilot Draft: NFSE-STUR)* was compiled (See APPENXID C).

Piloting the Draft Instrument

Following expert review, further feedback on the survey prompts was obtained by piloting the instrument. Lux (2010) suggests that a major purpose of piloting a draft instrument is face validity. And while face validity is not appropriate for later stages of item validation (DeVellis, 2003), requesting participants to comment on the items in this way is informative before application in a study (Netemeyer et al., 2003).

Participants in the pilot study were inservice teachers selected based on their enrollment in an online graduate level course about the new framework for science education at mid-sized northwestern university. Forty inservice teachers were invited to participate. Of these educators, thirteen completed the survey. Participants in the pilot study were K-5 teachers or 6-12 science teachers. Of those who elected to participate, only two did not have K-5 science teaching experience (one participant elected not to provide demographic data). The graduate course in which these participants were enrolled was focused on the first theme of the survey, Science & Engineering Practices. Most had also been through other professional development in science education and/or belonged to state or national science organizations.

The pilot survey (*Pilot Draft: NFSE-STUR*) began with written directions as well as an instructional video designed to facilitate participant success in completing the survey. As presented in Figure 2 and Figure 3 (see page 76), pilot survey participants were asked to complete the survey and provided the same framework quotes offered during the validation study. These participants, though, were also encouraged to provide comments about their experience using the instrument. This was done through a comment box placed the bottom of each page of the survey requesting participants to comment on prompts that they found confusing or difficult to answer for any reason. Statistical analysis on the internal consistency for the two sets of data, *understanding* and *readiness*, was conducted to establish the reliability of the instrument. Also, qualitative data collected from the open ended comment boxes were used to further refine the items in the pool before constructing the validation version of the instrument designed to assess

elementary teachers *understanding* and *readiness* to implement the new framework for science education. Based on its intended purpose, it was named, *New Framework for Science Education: Survey of Teacher Understanding & Readiness (Validation Draft: NFSE-STUR)*.

Administering the NFSE-STUR

Once the pilot testing the instrument concluded, a validation version of the instrument, *New Framework for Science Education: Survey of Teacher Understanding & Readiness (Validation Draft: NFSE-STUR)* (See APPENDIX D), was constructed and distributed, via email and listservs, to numerous K-5 classroom teachers in the states of Montana, Wyoming, Utah, and Idaho. The neighboring states of Wyoming, Utah, and Idaho were identified and selected for inclusion in the study based on similarities in political push-back in adopting the NGSS. Like Montana two of the three had officially put plans for adoption on hold, but all were continuing to conduct professional development regarding The Framework. The states of Washington and Oregon were eliminated from consideration since they had already adopted and were about to adopt, respectively, the NGSS.

Science educators and state-level policy makers facilitated the distribution of the survey instrument. Since the instrument was created online using Suveygizmo.com, (SurveyGizmo, 2014) participants accessed it through a link. As suggested by Nunnally (1978) as adequate for limiting associated participant error, $N=300$ participating teachers was the goal of this portion of the study (DeVellis, 2003). While data was collected for

teachers from K-12, only those participants with K-5 science teaching experience were included in this validation study.

Data Analysis Methods

Data from the process of the creation and implementation of the *NFSE-STUR* was used to evaluate the survey in terms of its validity and reliability. Following standards found in the literature (i.e. American Educational Research Association, 1999; DeVellis, 2003; Netemeyer et al., 2003) and examples of scale development (Bangert, 2004, 2006; Lux, 2010) validity of the survey was established. This involved both qualitative and quantitative measures: themes based upon the literature, an item pool developed from the operationalized themes, an expert review, and a pilot study all preceded the final validation study.

Though exploratory analysis with principal component analysis (PCA) and Exploratory Factor Analysis (EFA) are often used interchangeably (DeVellis, 2003), it is important to distinguish between the two since factor analysis demands the researcher to make a number of decisions about how the analysis is conducted—this is not true for most statistical analyses (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Principal component analysis is used to identify the fewest and most economical factors in a set of data (Netemeyer et al., 2003). PCA then is primarily about data reduction (Fabrigar et al., 1999; Netemeyer et al., 2003). The resulting principle components or factors are linear representations of original variables; therefore, they are grounded in the actual data (DeVellis, 2003). These characteristics make PCA the common and preferred factor extraction method in instrument development (McCoach et al., 2013). Exploratory

Factor Analysis (EFA) or Common Factor Analysis has a largely different purpose. It identifies the latent variables or underlying constructs that describe a set of items (Fabrigar et al., 1999; Netemeyer et al., 2003). To do this, it uses hypothetical variables instead of the actual data (DeVellis, 2003).

PCA and EFA often return similar results especially when communalities are high, PCA is a far simpler procedure (Fabrigar et al., 1999), PCA is the most common extraction method (McCoach et al., 2013), and PCA has been used with success in combination with CFA (Bangert, 2009). Because of these considerations, exploratory analysis with principal component analysis was conducted on the scales in this study.

Exploratory analysis and confirmatory factor analysis (CFA) were used to verify the underlying constructs of *NFSE-STUR* items written to measure elementary teachers' *understanding* of and *readiness* to implement The Framework. Exploratory analysis was used to determine the most economical factor model, and CFA was used to establish the goodness of fit for the hypothesized factor model produced by the exploratory analysis (Netemeyer et al., 2003).

Reliability of the *NFSE-STUR* instrument was established by measuring internal consistency using Cronbach's (1951) coefficient alpha. Considering the challenges of establishing reliability through other means (Netemeyer et al., 2003), the use of Cronbach's alpha is a common and satisfactory choice (DeVellis, 2003; Netemeyer et al., 2003).

CHAPTER FOUR

RESULTS

Introduction

In this chapter, the results from the study are presented. It was determined that a number of data were needed to answer the research questions posed: (1) What themes representing the implementation of The Framework are identified by the science education literature? (2) What are the underlying dimensions of *NFSE-STUR* items written to assess elementary teachers' *understanding* of and *readiness* to implement The Framework? Here, data and analyses are described for the expert review, the pilot study, and the final validation study.

Expert Review

During the expert review both qualitative and quantitative data were collected and analyzed in two stages. Expert participants rated each instrument item's relevance, clarity, and conciseness (DeVellis, 2003) on a four-point scale (4 = Very Strong, 3 = Strong, 2 = Weak, 1 = Very Weak). Descriptive statistics (see Table 1) were used to identify items considered for editing. Since Likert scale data is widely accepted as ordinal instead of interval data (Göb, McCollin, & Ramalhoto, 2007), medians instead of means were used for interpretation. All items with a median score of 3 or less on at least two areas were marked for further consideration. Reviewer comments (See APPENDIX E) became very

helpful in interpreting lower marks or raising questions to be address during stage two of the expert review.

Few items were changed based on data from stage one alone. Instead, concerns expressed in stage one of the review were used to leverage further expert insight in stage two. An example of this can be seen in items 5 and 6. Both received median scores of 3 or less in at least one category. In addition, multiple reviewers discussed each in their comments: “The use of models is very open ended—I think multiple questions might be needed here—to identify different types of models” and “This use of modeling for an engineering purpose is sufficiently different from its use in science, that I can see having it as a different prompt. That said, the necessity of a teacher understanding this is not as great as it is for the idea of modeling in a science context.” Challenges and concerns over prompts like this were addressed during the stage two focus group portion of the expert review.

Data from the expert review demonstrated the need for modifications in most items. Table 2 displays the original prompts alongside the edited prompts. In general, reviewer concerns and ideas were implemented in the new version of the instrument. Many of the items only required minor additions such as underlining key words. Some were refined to provide more clarity as was the case for item 15 where the words, “both word and graphic sources” were rewritten as “words, images, and other media.” There was also elimination of items by either merger or removal. The merging of items mostly involved the issue of dividing or uniting science and engineering. The first two prompts on the instrument are a good example of this procedure to combine related items.

Table 1. Descriptive Statistics for Expert Review of Initial Instrument Items

Prompt		Median	SD
1. When planning and teaching, educators have students participate in practices used by scientists in the real world.	Relevance	4	0
	Clarity	3	0.45
	Conciseness	4	0.54
2. When planning and teaching, educators have students participate in practices used by engineers in the real world.	Relevance	4	0.89
	Clarity	3	1.00
	Conciseness	4	1.30
3. When planning and teaching, educators have students ask questions about scientific phenomena that can drive exploration.	Relevance	4	0.00
	Clarity	4	0.89
	Conciseness	4	0.45
4. When planning and teaching, educators have students ask questions to define engineering problems that can drive design.	Relevance	4	0.45
	Clarity	3	0.84
	Conciseness	4	0.45
5. When planning and teaching, educators have students develop and use models (i.e. mental, conceptual, mathematical, or computer) to explain phenomena.	Relevance	4	0.00
	Clarity	2.5	0.96
	Conciseness	3.5	0.96
6. When planning and teaching, educators have students use models to analyze and test existing engineering systems or possible solutions.	Relevance	4	0.55
	Clarity	3	0.71
	Conciseness	3	0.55
7. When planning and teaching, educators have students plan and carry out investigations appropriate to the questions being asked.	Relevance	4	0.00
	Clarity	3	0.84
	Conciseness	4	0.55
8. When planning and teaching, educators have students plan and carry out investigations appropriate to the problem to be solved.	Relevance	4	1.00
	Clarity	3	0.55
	Conciseness	4	0.45
9. When planning and teaching, educators have students organize, analyze, and interpret observational data to make meaning.	Relevance	4	0.45
	Clarity	4	0.89
	Conciseness	4	0.55

Table 1 Continued

10. When planning and teaching, educators have students use mathematical thinking and computational skills to describe and predict phenomena.	Relevance	4	0.89
	Clarity	2.5	0.58
	Conciseness	3	0.82
11. When planning and teaching, educators have students use mathematical thinking and computational skills to inform engineering design.	Relevance	4	1.30
	Clarity	3	0.71
	Conciseness	3	0.55
12. When planning and teaching, educators have students construct evidence based explanations to describe phenomena that incorporate their understandings about science.	Relevance	4	0.00
	Clarity	3.5	0.96
	Conciseness	3.5	0.58
13. When planning and teaching, educators have students design and refine solutions that meet the needs of an engineering problem.	Relevance	4	0.82
	Clarity	3	0.50
	Conciseness	4	1.10
14. When planning and teaching, educators have students participate in discourse and argumentation—using evidence to make their case for either a scientific explanation or an engineering design.	Relevance	4	0.45
	Clarity	3	0.71
	Conciseness	3	0.55
15. When planning and teaching, educators have students obtain and evaluate information from both word and graphic sources.	Relevance	4	0.55
	Clarity	2	0.55
	Conciseness	3	0.55
16. When planning and teaching, educators have students communicate ideas clearly and persuasively through both words and graphics.	Relevance	4	0.89
	Clarity	2	0.55
	Conciseness	4	0.89
17. When planning and teaching, educators use concepts that bridge disciplinary (i.e. physical science, life science, or earth science) boundaries.	Relevance	4	1.34
	Clarity	2	1.14
	Conciseness	4	0.55
18. When planning and teaching, educators have students look for patterns as a strategy in observation or classification.	Relevance	4	0.00
	Clarity	3	0.45
	Conciseness	4	0.45

Table 1 Continued

19. When planning and teaching, educators have students discuss phenomena and develop explanation or designs in terms of cause and effect.	Relevance	4	0.00
	Clarity	3	0.84
	Conciseness	4	0.55
20. When planning and teaching, educators have students develop their understanding of the relationship of scale, proportion, and quantity in the macro and micro world.	Relevance	4	0
	Clarity	2	0.89
	Conciseness	3	0.84
21. When planning and teaching, educators have students view portions of the natural world as systems to understand interaction and interdependence.	Relevance	4	0.00
	Clarity	2	0.89
	Conciseness	3	0.84
22. When planning and teaching, educators have students consider the flow and conservation of energy and matter throughout the science curriculum.	Relevance	3	0.00
	Clarity	3	1.00
	Conciseness	3	0.55
23. When planning and teaching, educators have students investigate phenomena in terms of structure and function as a means of sense making.	Relevance	4	0.00
	Clarity	3	0.84
	Conciseness	4	0.55
24. When planning and teaching, educators have students identify things under investigation that remain stable and things that change	Relevance	4	0.00
	Clarity	3	0.84
	Conciseness	3	0.55
25. When planning and teaching, educators focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.	Relevance	4	0.00
	Clarity	4	1.10
	Conciseness	4	0.55
26. When planning and teaching, educators recognize that the development of student understanding of disciplinary core ideas is a progression that takes place over years.	Relevance	4	0.89
	Clarity	2	1.34
	Conciseness	4	0.89
27. When planning and teaching, educators provide opportunities for students to develop and understanding of core ideas in the physical sciences.	Relevance	4	1.41
	Clarity	2	1.10
	Conciseness	3	0.45
28. When planning and teaching, educators provide opportunities for students to develop and understanding of core ideas in the life sciences.	Relevance	4	1.41
	Clarity	2	1.10
	Conciseness	4	0.45

Table 1 Continued

29. When planning and teaching, educators provide opportunities for students to develop and understanding of core ideas in the earth and space sciences.	Relevance	4	1.41
	Clarity	2	1.10
	Conciseness	4	0.45
30. When planning and teaching, educators provide opportunities for students to develop and understanding of core ideas in engineering and the applications of science.	Relevance	4	1.41
	Clarity	2	0.45
	Conciseness	4	0.84
31. When planning and teaching, educators weave practices, concepts, and core ideas together into unified experiences.	Relevance	4	1.34
	Clarity	3	0.84
	Conciseness	4	0.00
32. When planning and teaching, educators have students explore disciplinary ideas by engaging in practices and making connections through crosscutting concepts.	Relevance	4	0.00
	Clarity	4	1.10
	Conciseness	4	0.45
33. When planning and teaching, educators intentionally select practices and concepts that best facilitate student sense making for particular core ideas.	Relevance	4	1.30
	Clarity	2	0.89
	Conciseness	4	0.45
34. When planning and teaching, educators have students use the crosscutting concepts when engaging in practices about disciplinary core ideas.	Relevance	4	1.34
	Clarity	4	1.10
	Conciseness	4	0.45
35. When planning and teaching, educators have students apply science and engineering practices within the context of core ideas.	Relevance	4	1.34
	Clarity	2	0.89
	Conciseness	4	0.45
36. When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.	Relevance	4	0.89
	Clarity	3	0.84
	Conciseness	4	0.55
37. When planning and teaching, educators develop scientific proficiency by having students participate in a number of scientific activities and thinking.	Relevance	4	1.41
	Clarity	2	1.10
	Conciseness	4	0.55

Table 1 Continued

38. When planning and teaching, educators use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.	Relevance	4	1.10
	Clarity	3	0.84
	Conciseness	4	0.45
39. When planning and teaching, educators have students engage in sustained investigations accompanied by necessary teacher support.	Relevance	4	0.89
	Clarity	3	1.00
	Conciseness	4	0.45
40. When planning and teaching, educators teach students how to present their scientific ideas and engineering solutions with clarity through both the written and spoken word.	Relevance	4	0.50
	Clarity	3.5	0.96
	Conciseness	4	0.00
41. When planning and teaching, educators teach students how mathematical concepts and skills apply to scientific investigations and engineering design.	Relevance	4	1.34
	Clarity	3	0.84
	Conciseness	4	0.55
42. When planning and teaching, educators teach students content area reading skills that facilitate their ability to access, interpret, and evaluate scientific knowledge.	Relevance	4	1.30
	Clarity	3	0.84
	Conciseness	3	0.55

In the original items, science and engineering were split—forming two separate items. Based on expert feedback, these were combined into one new item.

Only one area of the instrument experienced major changes. Original items 27-30 were removed and five new items were written. The structure and content of these items was not satisfactory—based on a median clarity score of 2 (See Table 1). One reviewer attached this comment to each of the items during stage one of the review, “I don’t think there is much to be gained asking whether it’s important to learn each discipline.” Even though other reviewers did not initially indicate a similar concern, as evidenced by the median relevance score of 4 for each item, the researcher brought up the issue during the stage two focus group. During that discussion all reviewers readily agreed with the benefits of rewriting the prompts. It was suggested that this aspect of The Framework was best captured by the NRC’s (2012) description of the criteria for a disciplinary core idea (See description on page 48 of this document).

All reviewer ideas and concerns were considered, and while most were used to make modifications to the instrument items, there were a few areas where reviewer suggestions were not followed. Through discussions with dissertation committee members, the researcher determined that there were places where reviewer suggestions did not best meet the intention of the instrument or went against principles found in the literature on instrument development. An example of this was the recommendation by one reviewer to change the structure of the prompt from “When planning and teaching, educators . . .” to “When planning and teaching, you . . .” In explaining this benefit, the reviewer said, “If you change the prompt to YOU above, it should be changed

everywhere. This makes it more personal. You should think about whether or not you want what teachers ARE doing or what they SHOULD be doing. I'm thinking you want to know they are doing." This suggestion was excellent for an instrument of alternative purpose, yet for the *NFSE-STUR* it was not—demonstrating a lack of adequate communication on the part of the researcher. More clarity on the purpose of the instrument should have been provided to the expert reviewers. Since the purpose of the instrument was to assess teacher *understanding* of and *readiness* to implement The Framework, which is new and includes conceptual shifts in science education, there was no expectation that teachers would already be practicing in these ways. The prompt structure was designed intentionally as a way of removing any personal expectation and therefore removing personal attack.

There were several draft items where reviewer suggestions were disregarded because it was determined that the original interpretation of The Framework was more accurate. Reviewers responded to items 31-35 by saying, "Question 32 covers all the ground that is needed for this section." "I am having a hard time differentiating between 31 and 32. This could be me." And, "This is a tough question -- I like it, but wonder if a second could be added about developing practices (explicitly) throughout the year." Through discussion with committee members and careful consideration of the NRC's (2012) document it was decided to leave these items unchanged for the pilot study. The expert review provided rich data that was used to dynamically change the original prompts. These changes produced a pool of items that clearly and concisely presented relevant aspects of The Framework.

Table 2. Item Changes Resulting from the Expert Review

Original Expert Review Item	Change	Resulting Pilot Instrument Item
1. When planning and teaching, educators have students participate in practices used by scientists in the real world.	Merged with #2 and key words underlined	When planning and teaching, educators have students participate in practices used by <u>scientists</u> and <u>engineers</u> in the real world.
2. When planning and teaching, educators have students participate in practices used by engineers in the real world.	Merged with #1 and key words underlined	
3. When planning and teaching, educators have students ask questions about scientific phenomena that can drive exploration.	Key words underlined	When planning and teaching, educators have students ask questions about <u>scientific phenomena</u> that can drive exploration.
4. When planning and teaching, educators have students ask questions to define engineering problems that can drive design.	Key words underlined	When planning and teaching, educators have students ask questions to define <u>engineering problems</u> that can drive design.
5. When planning and teaching, educators have students develop and use models (i.e. mental, conceptual, mathematical, or computer) to explain phenomena.	Reworded, split into two, and key words underlined	When planning and teaching, educators have students develop and refine <u>mental models</u> as part of the sense making process. When planning and teaching, educators have students develop and refine <u>conceptual models</u> to express their understanding about scientific phenomena.
6. When planning and teaching, educators have students use models to analyze and test existing engineering systems or possible solutions.	Reworded and key words underlined	When planning and teaching, educators have students develop models to visualize and refine an <u>engineered design</u> .

Table 2 Continued

7. When planning and teaching, educators have students plan and carry out investigations appropriate to the questions being asked.	Reworded, merged with #8, and key words underlined	When planning and teaching, educators have students plan and carry out investigations to gather data about <u>scientific phenomena</u> and <u>engineering problems</u> .
8. When planning and teaching, educators have students plan and carry out investigations appropriate to the problem to be solved.	Merged with #7	
9. When planning and teaching, educators have students organize, analyze, and interpret observational data to make meaning.	Unchanged	When planning and teaching, educators have students organize, analyze, and interpret observational data to make meaning.
10. When planning and teaching, educators have students use mathematical thinking and computational skills to describe and predict phenomena.	Reworded, merged with #11, and key words underlined	When planning and teaching, educators have students use mathematical thinking and computational skills to investigate <u>scientific questions</u> and <u>engineering problems</u> .
11. When planning and teaching, educators have students use mathematical thinking and computational skills to inform engineering design.	Merged with #10	
12. When planning and teaching, educators have students construct evidence-based explanations to describe phenomena that incorporate their understandings about science.	Key word underlined	When planning and teaching, educators have students construct evidence-based explanations to describe phenomena that incorporate their understandings about <u>science</u> .
13. When planning and teaching, educators have students design and refine solutions that meet the needs of an engineering problem.	Key word underlined	When planning and teaching, educators have students design and refine solutions that meet the needs of an <u>engineering</u> problem.

Table 2 Continued

<p>14. When planning and teaching, educators have students participate in discourse and argumentation—using evidence to make their case for either a scientific explanation or an engineering design.</p>	<p>Reworded and key words underlined</p>	<p>When planning and teaching, educators have students engage in evidence-based argumentation about <u>scientific explanations</u> or an <u>engineering designs</u>.</p>
<p>15. When planning and teaching, educators have students obtain and evaluate information from both word and graphic sources.</p>	<p>Reworded</p>	<p>When planning and teaching, educators have students obtain and evaluate information from words, images, and other media.</p>
<p>16. When planning and teaching, educators have students communicate ideas clearly and persuasively through both words and graphics.</p>	<p>Reworded</p>	<p>When planning and teaching, educators have students communicate ideas clearly and persuasively through words, images, and other media.</p>
<p>17. When planning and teaching, educators use concepts that bridge disciplinary (i.e. physical science, life science, or earth science) boundaries.</p>	<p>Unchanged</p>	<p>When planning and teaching, educators use concepts that bridge disciplinary (i.e. physical science, life science, or earth science) boundaries.</p>
<p>18. When planning and teaching, educators have students look for patterns as a strategy in observation or classification.</p>	<p>Unchanged</p>	<p>When planning and teaching, educators have students look for patterns as a strategy in observation or classification.</p>
<p>19. When planning and teaching, educators have students discuss phenomena and develop explanation or designs in terms of cause and effect.</p>	<p>Reworded and key words underlined</p>	<p>When planning and teaching, educators have students consider issues of cause and effect when questioning and discussing <u>scientific phenomena</u> or <u>engineering designs</u>.</p>
<p>20. When planning and teaching, educators have students develop their understanding of the relationship of scale, proportion, and quantity in the macro and micro world.</p>	<p>Reworded</p>	<p>When planning and teaching, educators have students develop an understanding that phenomena work differently at different scales.</p>

Table 2 Continued

<p>21. When planning and teaching, educators have students view portions of the natural world as systems to understand interaction and interdependence.</p>	<p>Reworded and key words underlined</p>	<p>When planning and teaching, educators have students use systems thinking when investigating <u>scientific phenomena</u>.</p>
<p>22. When planning and teaching, educators have students consider the flow and conservation of energy and matter throughout the science curriculum.</p>	<p>Reworded</p>	<p>When planning and teaching, educators have students consider that since energy and matter are conserved, much can be determined by studying their flow into and out of systems.</p>
<p>23. When planning and teaching, educators have students investigate phenomena in terms of structure and function as a means of sense making.</p>	<p>Unchanged</p>	<p>When planning and teaching, educators have students investigate phenomena in terms of structure and function as a means of sense making.</p>
<p>24. When planning and teaching, educators have students identify things under investigation that remain stable and things that change</p>	<p>Reworded</p>	<p>When planning and teaching, educators have students identify what aspects of a system remain stable over time and what aspects undergo patterns of change.</p>
<p>25. When planning and teaching, educators focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.</p>	<p>Unchanged</p>	<p>When planning and teaching, educators focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.</p>
<p>26. When planning and teaching, educators recognize that the development of student understanding of disciplinary core ideas is a progression that takes place over years.</p>	<p>Unchanged</p>	<p>When planning and teaching, educators recognize that the development of student understanding of disciplinary core ideas is a progression that takes place over years.</p>

Table 2 Continued

27. When planning and teaching, educators provide opportunities for students to develop and understanding of core ideas in the physical sciences.	Items 27 to 30 were removed and five new items were written	When planning and teaching, educators use a curriculum that follows a progression so that students engage in learning experiences that build on what they have previously learned and that prepare them for what they need to learn later.
28. When planning and teaching, educators provide opportunities for students to develop and understanding of core ideas in the life sciences.		When planning and teaching, educators include core ideas that have broad importance across multiple disciplines or are key organizing principles within a discipline.
29. When planning and teaching, educators provide opportunities for students to develop and understanding of core ideas in the earth and space sciences.		When planning and teaching, educators include core ideas that are important in investigating more complex ideas and solving problems.
30. When planning and teaching, educators provide opportunities for students to develop and understanding of core ideas in engineering and the applications of science.		When planning and teaching, educators include core ideas that relate to the interests and life experiences of students or societal concerns.
31. When planning and teaching, educators weave practices, concepts, and core ideas together into unified experiences.	Unchanged	When planning and teaching, educators weave practices, concepts, and core ideas together into unified experiences.

Table 2 Continued

<p>32. When planning and teaching, educators have students explore disciplinary ideas by engaging in practices and making connections through crosscutting concepts.</p>	<p>Unchanged</p>	<p>When planning and teaching, educators have students explore disciplinary ideas by engaging in practices and making connections through crosscutting concepts.</p>
<p>33. When planning and teaching, educators intentionally select practices and concepts that best facilitate student sense making for particular core ideas.</p>	<p>Unchanged</p>	<p>When planning and teaching, educators intentionally select practices and concepts that best facilitate student sense making for particular core ideas.</p>
<p>34. When planning and teaching, educators have students use the crosscutting concepts when engaging in practices about disciplinary core ideas.</p>	<p>Unchanged</p>	<p>When planning and teaching, educators have students use the crosscutting concepts when engaging in practices about disciplinary core ideas.</p>
<p>35. When planning and teaching, educators have students apply science and engineering practices within the context of core ideas.</p>	<p>Unchanged</p>	<p>When planning and teaching, educators have students apply science and engineering practices within the context of core ideas.</p>
<p>36. When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.</p>	<p>Unchanged</p>	<p>When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.</p>
<p>37. When planning and teaching, educators develop scientific proficiency by having students participate in a number of scientific activities and thinking.</p>	<p>Unchanged</p>	<p>When planning and teaching, educators develop scientific proficiency by having students participate in a number of scientific activities and thinking.</p>
<p>38. When planning and teaching, educators use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.</p>	<p>Unchanged</p>	<p>When planning and teaching, educators use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.</p>

Table 2 Continued

39. When planning and teaching, educators have students engage in sustained investigations accompanied by necessary teacher support.	Unchanged	When planning and teaching, educators have students engage in sustained investigations accompanied by necessary teacher support.
40. When planning and teaching, educators teach students how to present their scientific ideas and engineering solutions with clarity through both the written and spoken word.	Unchanged	When planning and teaching, educators teach students how to present their scientific ideas and engineering solutions with clarity through both the written and spoken word.
41. When planning and teaching, educators teach students how mathematical concepts and skills apply to scientific investigations and engineering design.	Unchanged	When planning and teaching, educators teach students how mathematical concepts and skills apply to scientific investigations and engineering design.
42. When planning and teaching, educators teach students content area reading skills that facilitate their ability to access, interpret, and evaluate scientific knowledge.	Unchanged	When planning and teaching, educators teach students content area reading skills that facilitate their ability to access, interpret, and evaluate scientific knowledge.

These revised items were compiled to form the *Pilot Draft: NFSE-STUR* which was used during the pilot stage of this instrument validation study.

Pilot Study

During the pilot study stage of the validation study, both quantitative and qualitative data were collected. The pilot draft of the instrument, *New Framework of Science Education Survey of Teacher Understanding & Readiness (Pilot Draft: NFSE-STUR)* was constructed based on the revised item pool created through expert review. This survey was administered to forty inservice teachers enrolled in an online graduate level course about the new framework for science education. Table 3 presents demographic data for the thirteen educators who chose to participate in the piloting of the instrument.

Table 3. Demographic Characteristics of Pilot Study Participants, $N = 13$

	<i>n</i>	Percent
Years of K-5 teaching experience		
0 – 5 years	4	33
6 – 10 years	4	33
11 – 15 years	1	8
16 – 20 years	0	0
20+ years	3	25
Hours of science instruction each week		
0 – 3 hours	7	58
4 – 6 hours	2	17
7+ hours	3	25
Rate your enjoyment of teaching science		
No Enjoyment	0	0
Slight Enjoyment	1	8
Fair Enjoyment	1	8
Strong Enjoyment	10	83

Table 3 Continued

Rate your success in teaching science		
No Success	0	0
Slight Success	1	8
Fair Success	6	50
Strong Success	5	42
Rate your Familiarity with the Current Best Practices in Science Education		
No Familiarity	0	0
Slight Familiarity	0	0
Fair Familiarity	6	50
Strong Familiarity	6	50
Involvement in Science Education		
Membership in State Science Organizations		
	1	8
Membership in National Science Organizations		
	8	67
Recent PD (beyond current inquiry class)		
	7	58

Note. Only 12 of the 13 participants are represented here since one chose not to provide demographic data.

The *Pilot Draft: NFSE-STUR* asked participants to rate their *understanding* and *readiness* to implement ideas from The Framework on similar four point scales (see Figures 1 and 2, page 76). At the bottom of each page of the online instrument, they were asked to provide comments about the items on the page and specifically encouraged to address anything confusing. Since establishing reliability is an important step in building a case for the validity of an instrument (DeVellis, 2003), internal consistency of the instrument items was determined by calculating Cronbach's Alpha (Cronbach, 1951) which is the most commonly used measure (Netemeyer et al., 2003). The analysis was conducted for items assessing each the *understanding* and *readiness* scales. The internal consistency of the *understanding* scale for the pilot study was, $\alpha = 0.981$. The internal consistency of the *readiness* scale for the pilot study was, $\alpha = 0.980$. These results, far

above the acceptable, $\alpha = 0.70$ (DeVellis, 2003), demonstrated the strength of the items association with each participant's true-score (Netemeyer et al., 2003).

Pilot participants submitted very few comments regarding confusion over items in the survey. This revealed the thoroughness of the expert review and presented strong face validity, an important aspect of piloting an instrument (Lux, 2010). Therefore, only one item was changed following the pilot study. Item 25 of the *Pilot Draft: NFSE-STUR* was a newly written item from the expert review. One participant commented, "Prompt 25 seems a little wordy. I had to reread it to make sure I was answering it correctly." Even though none of the other pilot participants raised a concern, the researcher sought the advice of two experts from the review panel to obtain suggestions about ways to phrase the item. The end product was far more clear and concise. Initially the item read, "When planning and teaching, educators use a curriculum that follows a progression so that students engage in learning experiences that build on what they have previously learned and that prepare them for what they need to learn late." After revision, the final version read, "When planning and teaching, educators use a learning progression approach by building from prior knowledge and working towards future sophistication."

The only other change in the instrument items from the pilot to final version was the removal of all underlining. Though no pilot participants raised an issue with the underlined words, the researcher struggled in making decisions about which words to underline throughout the instrument. The expert panel had suggested this as a way to facilitate participant completion of the instrument; however, the researcher was unsure

how best to determine when underlining was needed. Eliciting feedback from dissertation committee members provided a clear answer.

- Although I think there is much value to underlining as the reviewers suggested, especially for those key terms, it also comes at a price. I think we'll just have to what you underline could, in some ways, influence results - what words or phrases one participant perceives as relevant and key in an item, might not be perceived by another participant as such (N. Lux, personal communication, December 18, 2013).
- I also agree that it can be important to consider the recommendations of the reviewers...BUT, you do not have to follow what they recommend, they are, in fact, recommendations. I looked at the first page and thought, "way too much underlining" (L. Kelting-Gibson, personal communication, December 18, 2013).
- I also have an issue with too much underlining/italics. If we expect people to be able to read...and they are teachers...then they should be able to pick out the main points in a question (J. Graves, personal communication, December 18, 2013).
- I would suggest that you remove all underlining for the reasons that have already been expressed. Yes, I realize you want to help the reader interpret the question, but that is really up to him/her to read it carefully (and to reread it, if need be in order to answer the question) (A. Ellsworth, personal communication, December 18, 2013).

Following committee suggestions, all underlining was removed from the validation draft of the survey, *Validation Draft: NFSE-STUR* (See APPENDIX D).

The last revision to the instrument was a decision to change the scale from a four-point to six-point scale. This decision was based on the desire to tease out more of the variance in participant responses. Moving to a five-point scale would allow for a neutral response something that should be avoided if possible (Clark & Watson, 1995). A six-point scale eliminated this problem and also allowed for a better representation of the variance in teacher responses. This solution was suggested by a professional evaluator who stated that,

People will not give an answer that is at the end of the scale, so having extreme endpoints gives them permission to respond at the next most extreme point. In other words, if there is no "no understanding" option present, people will be reluctant to answer "slight understanding" since that puts them at the bottom of the barrel. So even if nobody responds to your highest and lowest options, I think it's important to have them there as options, to draw people away from the middle and toward the ends of the spectrum (M. Coe, personal communication, December, 20, 2013).

After the final revisions were made, the researcher felt confident about the instrument's readiness for the validation study.

Validation Study

Introduction

The *NFSE-STUR* was developed to assess in-service teacher *understanding* of and *readiness* to implement a new framework for science education as defined by the NRC's (2012) *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. The final version used in the validation study asked participants to rate their *understanding* on a six-point scale ranging from *No Understanding* to *Advanced Understanding*. Then, rate their *readiness* on a six-point scale ranging from *No Readiness* to *Advanced Readiness*.

In this section data from the final validation study are presented and analyzed. Since it was determined that *understanding* and *readiness* are two different constructs, these data were separated and treated as two scales during the analyses. This section first presents the demographic characteristics of the participants and the descriptive analysis of participant responses. Then, the remainder of the analyses are split into two separate

parts. The first presents the *understanding* scale, and the second presents the *readiness* scale.

Participant Demographics

Teachers across the states of Montana, Wyoming, Idaho, and Utah were invited to participate in the final validation of the *NFSE-STUR*. Though only teachers with K-5 experience were used in the validation of the instrument, middle school and high school science teachers were invited to participate as well since state officials in Montana and Utah expressed interest in sharing this data.

Table 4 displays demographic characteristics of the educators who participated in the validation study. More than half of the participants were currently teaching in Utah, about a quarter in Montana, and slightly less than that in Idaho. Very few K-5 teachers from Wyoming participated in the study. While most teachers (61%) reported teaching less than four hours of science a week, they overwhelmingly enjoyed what they did teach—92% reported fair or strong enjoyment. With 84% also reporting fair to strong success in science teaching, it is no surprise that nearly three-quarters also rated their understanding of current best practice at fair to strong.

Table 4. Demographic Characteristics of Validation Study Participants, $N = 167$

	<i>n</i>	Percent
Current State		
Montana	38	23
Idaho	25	15
Utah	100	60
Wyoming	4	2
Years of K-5 teaching experience		
0 – 5 years	60	36
6 – 10 years	37	22

Table 4 Continued

11 – 15 years	23	14
16 – 20 years	18	11
20+ years	29	17
Hours of science instruction each week*		
Not reporting	3	2
0 – 3 hours	101	61
4 – 6 hours	44	26
7+ hours	19	11
Rate your enjoyment of teaching science		
Not reporting	2	1
No Enjoyment	1	0
Slight Enjoyment	11	7
Fair Enjoyment	56	34
Strong Enjoyment	97	58
Rate your success in teaching science		
Not reporting	4	2
No Success	4	2
Slight Success	20	12
Fair Success	90	54
Strong Success	49	30
Rate your District's Commitment to Science Education		
Not reporting	1	0
No Commitment	5	3
Slight Commitment	38	23
Fair Commitment	85	51
Strong Commitment	38	23
Rate your Familiarity with the Current Best Practices in Science Education		
Not reporting	1	0
No Familiarity	8	4
Slight Familiarity	47	28
Fair Familiarity	79	47
Strong Familiarity	35	21
Involvement in Science Education		
Membership in State Science Organizations		
	20	12
Membership in National Science Organizations		
	24	14
Inservice Science PD	123	74

Descriptive Analysis

To determine the normality of the data, descriptive analyses were conducted. Means and standard deviations for each instrument item are presented in Table 5. Since participant *understanding* and *readiness* to implement were both assessed and have been determined to measure separate constructs, they are displayed separately in the table.

Attempting to fit data to measurement models using exploratory analyses with nonnormal data can lead to error filled conclusions (Fabrigar et al., 1999). For this reason, descriptive analysis also examined each construct's normality. West, Finch, and Curran (1995) recommend a threshold of skew < 2 and kurtosis < 7 when attempting exploratory analysis. The skewedness and kurtosis for each whole construct was calculated and found to be well under the threshold. While four individual items for the *understanding* construct were skewed beyond the threshold, no items departed from normal "peakness." And the symmetry of the whole construct (skew = .061, SE = .188) and "peakness" (kurtosis = -.185, SE = .374) did not depart significantly from normality ($W = .992$, $p = .521$).

The symmetry of the *readiness* construct (skew = .179, SE = .188) and "peakness" (kurtosis = -.158, SE = .375) also did not depart significantly from normality ($W = .991$, $p = .418$) even though six individual items did exceed the symmetry threshold. None of the individual items for the *readiness* construct surpassed the "peakness" threshold.

Table 5. Descriptive Statistics for *Understanding* and *Readiness* Scales

	Understanding		Readiness	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1. When planning and teaching, educators have students participate in practices used by scientists and engineers in the real world.	3.82	1.08	3.50	1.23
2. When planning and teaching, educators have students ask questions about scientific phenomena that can drive exploration.	3.96	1.10	3.67	1.23
3. When planning and teaching, educators have students ask questions to define engineering problems that can drive design.	3.14	1.18	2.87	1.28
4. When planning and teaching, educators have students develop and refine mental models as part of the sense making process.	3.20	1.27	3.01	1.34
5. When planning and teaching, educators have students develop and refine conceptual models to express their understanding about scientific phenomena.	3.33	1.23	3.12	1.34
6. When planning and teaching, educators have students develop models to visualize and refine an engineered design.	3.39	1.26	3.07	1.40
7. When planning and teaching, educators have students plan and carry out investigations to gather data about scientific phenomena and engineering problems.	4.07	1.18	3.69	1.33
8. When planning and teaching, educators have students organize, analyze, and interpret observational data to make meaning.	4.08	1.05	3.81	1.22
9. When planning and teaching, educators have students use mathematical thinking and computational skills to investigate scientific questions and engineering problems.	3.67	1.11	3.32	1.30

Table 5 Continued

10. When planning and teaching, educators have students construct evidence-based explanations to describe phenomena that incorporate their understandings about science.	3.83	1.21	3.45	1.28
11. When planning and teaching, educators have students design and refine solutions that meet the needs of an engineering problem.	3.09	1.15	2.74	1.26
12. When planning and teaching, educators have students engage in evidence-based argumentation about scientific explanations or an engineering designs.	3.50	1.22	3.10	1.32
13. When planning and teaching, educators have students obtain and evaluate information from words, images, and other media.	4.14	1.07	3.89	1.25
14. When planning and teaching, educators have students communicate ideas clearly and persuasively through words, images, and other media.	4.06	1.05	3.73	1.19
15. When planning and teaching, educators use concepts that bridge disciplinary (i.e. physical science, life science, or earth science) boundaries.	3.86	1.12	3.56	1.28
16. When planning and teaching, educators have students look for patterns as a strategy in observation or classification.	4.11	1.06	3.84	1.21
17. When planning and teaching, educators have students consider issues of cause and effect when questioning and discussing scientific phenomena or engineering designs.	4.07	1.00	3.81	1.12
18. When planning and teaching, educators have students develop an understanding that phenomena work differently at different scales.	3.19	1.21	2.85	1.28
19. When planning and teaching, educators have students uses systems thinking when investigating scientific phenomena.	2.94	1.27	2.69	1.34
20. When planning and teaching, educators have students consider that since energy and matter are conserved, much can be determined by studying their flow into and out of systems.	3.14	1.27	2.70	1.33

Table 5 Continued

21. When planning and teaching, educators have students investigate phenomena in terms of structure and function as a means of sense making.	3.06	1.29	2.80	1.33
22. When planning and teaching, educators have students identify what aspects of a system remain stable over time and what aspects undergo patterns of change.	3.35	1.22	2.93	1.22
23. When planning and teaching, educators focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.	4.40	0.99	4.01	1.18
24. When planning and teaching, educators recognize that the development of student understanding of disciplinary core ideas is a progression that takes place over years.	4.50	1.00	4.10	1.23
25. When planning and teaching, educators use a learning progression approach by building from prior knowledge and working towards future sophistication.	4.44	1.02	4.10	1.20
26. When planning and teaching, educators include core ideas that have broad importance across multiple disciplines or are key organizing principles within a discipline.	4.14	1.06	3.76	1.21
27. When planning and teaching, educators include core ideas that are important in investigating more complex ideas and solving problems.	3.98	1.04	3.61	1.14
28. When planning and teaching, educators include core ideas that relate to the interests and life experiences of students or societal concerns.	4.27	1.02	3.88	1.24
29. When planning and teaching, educators include core ideas that can be investigated at multiple grade levels with increasing sophistication.	4.10	1.11	3.65	1.27
30. When planning and teaching, educators weave practices, concepts, and core ideas together into unified experiences.	3.99	1.06	3.61	1.19
31. When planning and teaching, educators have students explore disciplinary ideas by engaging in practices and making connections through crosscutting concepts.	3.57	1.26	3.13	1.30

Table 5 Continued

32. When planning and teaching, educators intentionally select practices and concepts that best facilitate student sense making for particular core ideas.	3.90	1.13	3.55	1.25
33. When planning and teaching, educators have students use the crosscutting concepts when engaging in practices about disciplinary core ideas.	3.28	1.26	2.95	1.31
34. When planning and teaching, educators have students apply science and engineering practices within the context of core ideas.	3.56	1.12	3.20	1.23
35. When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.	4.84	0.89	4.40	1.15
36. When planning and teaching, educators develop scientific proficiency by having students participate in a number of scientific activities and thinking.	4.67	0.89	4.17	1.19
37. When planning and teaching, educators use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.	4.49	1.00	3.90	1.26
38. When planning and teaching, educators have students engage in sustained investigations accompanied by necessary teacher support.	4.33	1.09	3.80	1.26
39. When planning and teaching, educators teach students how to present their scientific ideas and engineering solutions with clarity through both the written and spoken word.	4.16	1.03	3.59	1.23
40. When planning and teaching, educators teach students how mathematical concepts and skills apply to scientific investigations and engineering design.	3.83	1.13	3.36	1.25
41. When planning and teaching, educators teach students content area reading skills that facilitate their ability to access, interpret, and evaluate scientific knowledge.	4.41	0.99	3.99	1.15

Understanding Scale

Exploratory Analysis. Sample size in factor analysis is a much discussed issue, yet there exists limited agreement. Tinsley and Tinsley (1987) argue that while it is commonly assumed that “bigger is better,” sample sizes higher than 300 rarely change the stability of factor loading. The authors note that others argue the best approach to sample size is by using a ratio of instrument items to participants (i.e. 1:5 or even 1:10). DeVellis (2003), though, states that the ratio approach should not be used as a hard and fast rule since the needed ratio is not consistent—decreasing as the number of items increases.

Fabrigar et al. (1999) suggest that the ratio method is not sensitive to various characteristics in a set of data anyway. “The primary limitation of such guidelines is that adequate sample size is not a function of the number of measured variables per se but is instead influenced by the extent to which factors are overdetermined and the level of the communalities of the measured variables” (p. 274). Practically, it is not uncommon to see factor analysis used on modest samples of 150—especially when best practice suggests to split the sample and replicate the analysis on separate random samples to establish generalizability (DeVellis, 2003). Research has shown that stable loadings can be generated with ratios as small as 1:3 for instruments with as few as 20 items (Tinsley et al., 1987). The important considerations with small sample sizes is beginning with at least four items per assumed factor and having communalities that average .70 and higher (Fabrigar et al., 1999). With this literature as support, the $N = 167$ sample for the *NFSE-*

STUR was randomly split to conduct first the exploratory analysis with principle component analysis (PCA) and then the confirmatory factor analysis (CFA).

As recommended for scale development, exploratory analysis procedures using PCA were conducted using a random subsample ($n = 83$) of in-service teacher responses to the instrument. The resulting correlation matrix was evaluated for multicollinearity and items exhibiting extremely high correlations were removed. Items 4, 8, 16, and 36 were removed as they correlated with at least one other item higher than the suggest threshold of .80 (Field, 2013). With these items removed, the factorability of the correlation matrix, Kaiser-Meyer-Olkin Measure of Sampling Adequacy (.912) and Bartlett's Test of Sphericity ($\chi^2_{666} = 3,101.06, \rho < .001$), indicated that the data were appropriate for the analysis to proceed.

The clearest and most explainable factor model emerged using maximum likelihood extraction and Varimax rotation methods. A five-factor solution was determined using the best know procedure, the Kaiser criterion based on eigenvalues higher than 1.00 (Fabrigar et al., 1999) and was verified using the Scree Test as recommended by Cattell (1966). The rotated solution shown in Table 6, presented interpretable factors: Science & Engineering Practices, Teaching Disciplinary Core Ideas, Crosscutting Concepts, Integration of the Three Dimensions, and Best Practices in Science Education. These five factors contributed to explain over 74% of the total item variance. The first factor was responsible for contributing to over 54% of the variance.

Table 6. Rotated Factor Structure for the *NFSE-STUR Understanding Scale*

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Science & Engineering Practices					
3. When planning and teaching, educators have students ask questions to define engineering problems that can drive design.	.795	.049	.220	.359	-.014
2. When planning and teaching, educators have students ask questions about scientific phenomena that can drive exploration.	.747	.190	.170	.302	.026
5. When planning and teaching, educators have students develop and refine conceptual models to express their understanding about scientific phenomena.	.728	.193	.305	.194	-.001
11. When planning and teaching, educators have students design and refine solutions that meet the needs of an engineering problem.	.728	.043	.353	.278	.277
6. When planning and teaching, educators have students develop models to visualize and refine an engineered design.	.725	.190	.249	.156	.208
7. When planning and teaching, educators have students plan and carry out investigations to gather data about scientific phenomena and engineering problems.	.719	.484	.257	-.112	.102
14. When planning and teaching, educators have students communicate ideas clearly and persuasively through words, images, and other media.	.670	.489	.142	.027	.268
10. When planning and teaching, educators have students construct evidence-based explanations to describe phenomena that incorporate their understandings about science.	.688	.350	.406	.156	.064
1. When planning and teaching, educators have students participate in practices used by scientists and engineers in the real world.	.655	.257	.229	.213	.084
9. When planning and teaching, educators have students use mathematical thinking and computational skills to investigate scientific questions and engineering problems.	.622	.275	.293	.164	.267

Table 6 Continued

12. When planning and teaching, educators have students engage in evidence-based argumentation about scientific explanations or an engineering designs.	.600	.295	.484	.036	.244
Teaching Disciplinary Core Ideas					
25. When planning and teaching, educators use a learning progression approach by building from prior knowledge and working towards future sophistication.	1.71	.796	.118	.250	.270
35. When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.	.217	.769	.002	.099	.332
24. When planning and teaching, educators recognize that the development of student understanding of disciplinary core ideas is a progression that takes place over years.	.215	.737	.156	.251	.297
23. When planning and teaching, educators focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.	.316	.730	.281	.232	.020
27. When planning and teaching, educators include core ideas that are important in investigating more complex ideas and solving problems.	.267	.691	.353	.249	.125
28. When planning and teaching, educators include core ideas that relate to the interests and life experiences of students or societal concerns.	.284	.654	.200	.429	.204
26. When planning and teaching, educators include core ideas that have broad importance across multiple disciplines or are key organizing principles within a discipline.	.346	.606	.257	.416	.069
Crosscutting Concepts					
21. When planning and teaching, educators have students investigate phenomena in terms of structure and function as a means of sense making.	.257	.154	.792	.202	.201
18. When planning and teaching, educators have students develop an understanding that phenomena work differently at different scales.	.305	.149	.743	.243	.127

Table 6 Continued

19. When planning and teaching, educators have students use systems thinking when investigating scientific phenomena.	.333	.017	.722	.102	.201
20. When planning and teaching, educators have students consider that since energy and matter are conserved, much can be determined by studying their flow into and out of systems.	.325	.124	.698	.333	-.016
22. When planning and teaching, educators have students identify what aspects of a system remain stable over time and what aspects undergo patterns of change.	.290	.321	.692	.096	.172
17. When planning and teaching, educators have students consider issues of cause and effect when questioning and discussing scientific phenomena or engineering designs.	.385	.363	.614	.151	.175
Integration of the Three Dimensions					
33. When planning and teaching, educators have students use the crosscutting concepts when engaging in practices about disciplinary core ideas	.298	.281	.238	.723	.187
31. When planning and teaching, educators have students explore disciplinary ideas by engaging in practices and making connections through crosscutting concepts.	.216	.324	.249	.656	.328
32. When planning and teaching, educators intentionally select practices and concepts that best facilitate student sense making for particular core ideas.	.206	.353	.338	.585	.271
Best Practices in Science Education					
39. When planning and teaching, educators teach students how to present their scientific ideas and engineering solutions with clarity through both the written and spoken word.	.128	.358	.282	.162	.706
40. When planning and teaching, educators teach students how mathematical concepts and skills apply to scientific investigations and engineering design.	.128	.269	.458	.245	.626

Table 6 Continued

37. When planning and teaching, educators use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.	.224	.456	.078	.386	.577
38. When planning and teaching, educators have students engage in sustained investigations accompanied by necessary teacher support.	.294	.300	.218	.450	.571

The second factor contributed to almost 8%, the third factor over 5%, the fourth factor over 3% and the fifth factor almost 3% of the total item variance. Items that severely cross-loaded were removed.

- Item 13 loaded on factor 2 (.605) but cross-loaded with factor 1 (.528).
- Item 15 loaded on factor 3 (.483) but cross-loaded with factor 4 (.397).
- Item 29 loaded on factor 2 (.533) but cross-loaded on factor 4 (.478).
- Item 30 loaded on factor 2 (.510) but cross-loaded with factor 4 (.441).
- Item 34 loaded on factor four (.496) but cross-loaded with factor 3 (.422).
- Item 41 loaded on factor 2 (.620) but cross-loaded with factor 5 (.564).

These six items were removed. All other items were found to have factor loadings of .400 or greater (Stevens, 2002) and were retained for the final version of the *understanding* scale (See APPENDIX F).

Whole Scale and Factor Reliability Analysis. Cronbach's Coefficient Alphas were calculated for the entire *understanding* scale and for each of the five factors resulting from the exploratory analysis. Table 7 presents these results which demonstrate high levels of internal consistency and therefore reliability. Cronbach's Alpha is the most common measure used to determine internal consistency (Netemeyer et al., 2003). Alpha scores higher than .70 are considered acceptable (DeVellis, 2003). Even the lowest of the individual factor scores, Best Practices in Science Education $\alpha = .876$, was well above the acceptable standard.

Table 7. Whole Scale & Factor Reliability for *NFSE-STUR Understanding Construct*

Latent Variable	Survey Items from exploratory analysis	α
Entire <i>Understanding</i> Scale	All in Final Version	.971
Science & Engineering Practices	1-3, 5-7, 9-12, 14	.952
Teaching Disciplinary Core Ideas	23-28, 35	.942
Crosscutting Concepts	17-22	.915
Integration of the Three Dimensions	31-33	.883
Best Practices in Science Education	37-40	.876

Confirmatory Factor Analysis. While confirmatory factor analysis (CFA) is a primary tool in instrument design, there is not clear consensus about the indexes that are most helpful in evaluating a predicted model's fit (Bangert, 2006). MacCallum, Browne, and Sugawara (1996) suggest three measures that provide key information in interpreting the fit of a model: the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), and the Non-Normed Fit Index (NNFI). Others also agree with this approach; Bangert (2006) explains why. The RMSEA describes the amount of error there is between the predicted model and the estimated population. CFI and NNFI further describe the model fit by taking sample size and model complexity into account. RMSEAs less than .05 are a close fit and those between .05 and .08 are considered a fair fit (MacCallum et al., 1996). A good model fit also has CFI and NNFI values above .90 (Bangert, 2006).

After completing the exploratory analysis, CFA was conducted using Lisrel 8.72 (Joreskog & Sorbom, 2001) with the second random subsample ($n = 84$). The purpose of conducting a CFA is to confirm a predicted pattern or one made apparent by previous analytics (DeVellis, 2003). In this research, the CFA was used to confirm the factor pattern (latent variable model) that arose from the exploratory analysis with PCA.

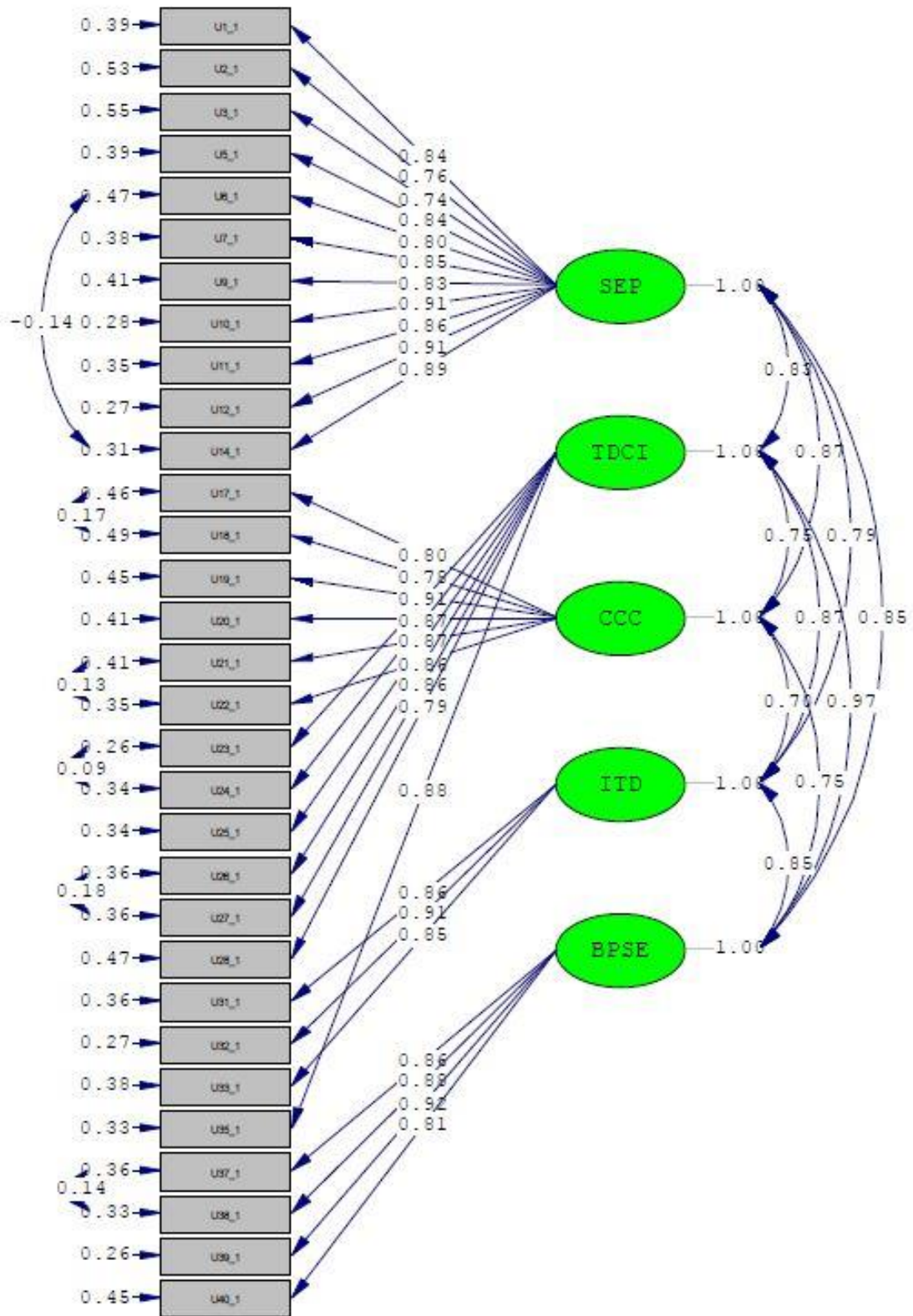


Figure 5. Understanding Scale Model Fit

Results from the CFA indicated that the independence model (which tests the hypothesis that all variables are uncorrelated) could be rejected, ($\chi^2_{465} = 12,836.49, \rho < .001$). The five-factor latent variable model that was developed from the exploratory analysis proved to be a superior fit to the subsample, ($\chi^2_{419} = 620.02, \rho < .001$). Figure 5 displays this model fit.

The five-factor model produced during the exploratory analysis produced a RMSEA of .065. The 90% confidence interval (.051 - .079) surrounding the RMSEA result provides supporting evidence that the proposed model is a fair fit to the estimated population. The accuracy of this fit is strengthened by a CFI of .98 and a NNFI of .98—both well above the suggested threshold. The power of this model fit was determined to exceed .90 when using Kim's (2005) tables which show the relationship between power, model degrees of freedom ($df = 419$) and the value of the model noncentrality parameter ($NCP = 149$) or the chi square distribution used to evaluate model fit.

Readiness Scale

Exploratory Analysis. Exploratory analysis procedures with principal component analysis were also conducted on the *readiness* scale using the same random subsample ($n = 83$) of in-service teacher responses to the instrument. The resulting correlation matrix was first evaluated for multicollinearity and items exhibiting extremely high correlations were removed. Items 4, 13, 27, 31, and 36 were removed as they correlated with at least one other item higher than the suggest threshold of .80 (Field, 2013). With these items removed, the factorability of the correlation matrix, Kaiser-Meyer-Olkin Measure of

Sampling Adequacy (.926) and Bartlett's Test of Sphericity ($\chi^2_{630} = 3,208.07, p < .001$), indicated that the data were appropriate for the analysis to proceed.

The clearest and most explainable factor model emerged using maximum likelihood extraction and Varimax rotation methods. A four-factor solution was determined using the best know procedure, the Kaiser criterion based on eigenvalues higher than 1.00 (Fabrigar et al., 1999) and was verified using the Scree Test as recommended by Cattell (1966). The rotated solution shown in Table 8, presented interpretable factors: Students Learning as Scientists and Engineers, Integration and Real-World Application of Core Ideas, Best Practices for Student Learning, and Crosscutting Concepts. These four factors contributed to explain over 74% of the total item variance. The first factor was responsible for contributing to over 61% of the variance. The second factor contributed to almost 6%, the third factor almost 4%, and the fourth factor over 3% of the total item variance. Items that severely cross-loaded were removed.

- Item 3 loaded on factor 3 (.465) but cross-loaded with factor 1 (.462).
- Item 16 loaded on factor 1 (.556) but cross-loaded with factor 3 (.519).

These two items were removed. All other items were found to have factor loadings of .400 or greater (Stevens, 2002) and were retained for the final version of the *readiness* scale (See APPENDIX G).

Table 8. Rotated Factor Structure for the *NFSE-STUR Readiness Scale*

	Factor 1	Factor 2	Factor 3	Factor 4
Students Learning as Scientists and Engineers				
10 When planning and teaching, educators have students construct evidence-based explanations to describe phenomena that incorporate their understandings about science.	.761	.397	.183	.208
8 When planning and teaching, educators have students organize, analyze, and interpret observational data to make meaning.	.697	.485	.238	.197
5 When planning and teaching, educators have students develop and refine conceptual models to express their understanding about scientific phenomena.	.672	.341	.318	.262
12 When planning and teaching, educators have students engage in evidence-based argumentation about scientific explanations and engineered designs.	.671	.213	.307	.470
1 When planning and teaching, educators have students participate in practices used by scientists and engineers in the real world.	.671	.142	.447	.274
7 When planning and teaching, educators have students plan and carry out investigations to gather data about scientific phenomena and engineering problems.	.644	.279	.245	.348
11 When planning and teaching, educators have students design and refine solutions that meet the needs of an engineering problem.	.636	.371	.178	.468
9 When planning and teaching, educators have students apply mathematical and computational thinking to investigate scientific questions and engineering problems.	.621	.512	.134	.251
17 When planning and teaching, educators have students consider issues of cause and effect when questioning and discussing scientific phenomena or engineering designs.	.608	.126	.491	.379
15 When planning and teaching, educators use concepts that bridge disciplinary (i.e. physical science, life science, or earth science) boundaries.	.605	.316	.370	.377
6 When planning and teaching, educators have students develop models to visualize and refine an engineered design.	.591	.302	.277	.456
14 When planning and teaching educators have students communicate ideas clearly and persuasively through words, images, and other media.	.572	.431	.345	.211

Table 8 Continued

Integration and Real-World Application of Core Ideas

32 When planning and teaching, educators intentionally select practices and concepts that best facilitate student sense making for particular core ideas.	.225	.743	.270	.218
30 When planning and teaching, educators weave practices, concepts, and core ideas together into unified experiences.	.296	.728	.307	.182
33 When planning and teaching, educators have students use the crosscutting concepts when engaging in practices about disciplinary core ideas.	.294	.726	.189	.260
34 When planning and teaching, educators have students apply science and engineering practices within the context of core ideas.	.319	.700	.288	.347
25 When planning and teaching, educators use a learning progression approach by building from prior knowledge and working towards future sophistication.	.350	.659	.453	.056
26 When planning and teaching, educators include core ideas that have broad importance across multiple disciplines or are key organizing principles within a discipline.	.405	.618	.454	.184
41 When planning and teaching, educators teach students content area reading skills that facilitate their ability to access, interpret, and evaluate scientific knowledge.	.073	.614	.460	.264
40 When planning and teaching, educators teach students how mathematical concepts and skills apply to scientific investigation and engineering design.	.297	.605	.211	.418
39 When planning and teaching, educators teach students how to present their scientific ideas and engineering solutions with clarity through both the written and spoken word.	.214	.570	.396	.377
28 When planning and teaching, educators include core ideas that relate to the interests and life experiences of students or societal concerns.	.352	.565	.414	.159
Best Practices for Student Learning				
35 When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.	.251	.386	.769	.164
24 When planning and teaching, educators recognize that the development of student understandings of disciplinary core ideas is a progression that takes place over years.	.326	.375	.759	.089

Table 8 Continued

29 When planning and teaching, educators include core ideas that can be investigated at multiple grade levels with increasing sophistication.	.302	.321	.718	.240
37 When planning and teaching, educators use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.	.163	.351	.707	.366
23 When planning and teaching, educators focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.	.357	.406	.672	.084
38 When planning and teaching, educators have students engage in sustained investigations accompanied by necessary teacher support.	.160	.415	.663	.389
2 When planning and teaching, educators have students ask questions about scientific phenomena that can drive exploration.	.426	.194	.634	.344
Crosscutting Concepts				
20 When planning and teaching, educators have students consider that since energy and matter are conserved, much can be determined by studying their flow into and out of systems.	.416	.166	.170	.748
19 When planning and teaching, educators have students use systems thinking when investigating scientific phenomena.	.349	.328	.231	.708
21 When planning and teaching, educators have students investigate phenomena in terms of structure and function as a means of sense making.	.248	.326	.278	.705
18 When planning and teaching, educators have students develop an understanding that phenomena work differently at different scales.	.506	.276	.188	.659
22 When planning and teaching, educators have students identify what aspects of a system remain stable over time and what aspects undergo patterns of change.	.364	.290	.370	.554

Whole Scale and Factor Reliability Analysis. Cronbach's Coefficient Alphas were calculated for the entire *readiness* scale and for each of the four factors resulting from the exploratory analysis. Table 9 presents these results which demonstrate high levels of internal consistency and therefore reliability. Cronbach's Alpha is the most common measure used to determine internal consistency (Netemeyer et al., 2003). Alpha scores higher than .70 are considered acceptable (DeVellis, 2003). Even the lowest of the individual factor scores, Crosscutting Concepts $\alpha = .917$ was well above the acceptable standard.

Table 9. Whole Scale and Factor Reliability for *NFSE-STUR Readiness* Construct

Latent Variable	Survey Items from exploratory analysis	α
Entire <i>Readiness</i> Scale	All in Final Version	.981
Students Learn as Scientists & Engineers	1, 5-12, 14-15, 17	.964
Integration and Real-World Application of Core Ideas	25-26, 28, 30, 32-34, 39-41	.950
Best Practices for Student Learning	2, 23-24, 29, 35, 37-38	.949
Challenging Crosscutting Concepts	18-22	.917

Confirmatory Factor Analysis. After completing the exploratory analysis with PCA, CFA was conducted using Lisrel 8.72 (Joreskog & Sorbom, 2001) with the second random subsample ($n = 84$). The purpose of conducting a CFA is to confirm a predicted pattern or one made apparent by previous analytics (DeVellis, 2003). In this research, the CFA was used to confirm the factor pattern (latent variable model) that arose from the exploratory analysis. Results from the CFA indicated that the independence model (which tests the hypothesis that all variables are uncorrelated) could be rejected, ($\chi^2_{561} = 15,539.09, \rho < .001$).

The four-factor latent variable model that was developed from the exploratory analysis proved to be a superior fit to the subsample, ($\chi^2_{515} = 663.36, \rho < .001$). Figure 6 displays this model fit.

The four-factor model developed from the exploratory analysis produced a RMSEA of .056. The 90% confidence interval (.041 - .069) surrounding the RMSEA result provides supporting evidence that the proposed model is a fair fit to the estimated population. The accuracy of this fit is strengthened by a CFI of .99 and a NNFI of .99—both well above the suggested threshold. The power of this model fit was determined to exceed .90 when using Kim's (2005) tables which show the relationship between power, model degrees of freedom ($df = 515$) and the value of the model noncentrality parameter ($NCP = 134$) for the chi square distribution used to evaluate model fit

Post Hoc Analysis on Crosscutting Concepts Theme

The splitting of the items written for the Crosscutting Concepts theme created the need for further analysis. On the *readiness* scale, the exploratory analysis divided these items into two separate factors. While the second factor was comprised entirely of items from the original theme (18 – 22), the first factor consisted mostly of the items from the original theme of Science & Engineering Practices. Yet, items 15 – 17, intended for the Crosscutting Concepts theme also loaded on this first factor. Describing this loading through qualitative measures alone proved difficult. To explore it further, the sum of the means for items 15 – 17 and 18 – 22 were calculated and compared using a paired-samples *t* test. G*Power 3.1.6 (Faul, Erdfelder, Lang, & Buchner, 2013) was used to determine the appropriate sample size needed to achieve a high power, Power = .95. A

random subsample ($n = 8$) was used to compare the collective means. The variance in the first three items, $m = 3.54, SD = 1.25$ and the next five items, $m = 2.13, SD = 1.08$ was found to be statistically significant, $t_7 = 4.88, p = .002$ with a very large effect, $d = 1.21$.

Chapter Summary

Data were gathered and analyzed during three different stages of the research. First, qualitative and quantitative data were gathered during the expert review. These data were analyzed using descriptive and qualitative measures to make validating modifications to the instrument. Next, during the pilot study, qualitative and quantitative data were gathered to further refine the instrument. The participant responses were analyzed for internal consistency using Cronbach's coefficient alpha. Finally quantitative data were gathered in a validation study effort. These data were analyzed descriptively to determine their normality prior to further analysis. Then, using exploratory analysis with principal component analysis, latent variables were determined and were found to be a good fit to the population model as identified using confirmatory factor analysis. Internal consistency of the scale and the individual factors were established using Cronbach's coefficient alpha.

CHAPTER FIVE

DISCUSSION

Introduction

Major shifts in science education are underway; however, it is currently unclear how successful elementary educators will be in implementing The Framework, a new framework for science education as described by, *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Research Council, 2012) and ultimately the new national standards, *Next Generation Science Standards: For States, By States* (NGSS Lead States, 2013). With no way to measure the level of readiness for implementation, it is difficult for state agencies, professional developers, and school districts to plan for professional development training. It leaves teacher preparation programs at a disadvantage in attempting to prepare preservice educators for the challenges that lie ahead.

The purpose of this study was to identify key constructs of The Framework and to develop a *New Framework for Science Education Survey of Teacher Understanding & Readiness (NFSE-STUR)* to assess elementary educators' *understanding* of and *readiness* for these constructs. The study was undertaken to identify the themes critical for implementing The Framework at the elementary level and to validate an instrument that can be used to assess elementary teachers' self-perceptions of science education related to their *understanding* of and *readiness* to successfully implement The Framework.

This study examined the NRC's (2012) report, *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*, and related documents to determine the key themes for successful implementation of The Framework. Using these elements, the *NFSE-STUR* was created and validated. Two questions were used to guide the research.

1. What themes representing the implementation of the new framework for science education are identified by the science education literature?
2. What are the underlying dimensions of *NFSE-STUR* items written to assess elementary teachers' *understanding* of and *readiness* to implement The Framework?

In this chapter, the results and implications of the study are discussed. First, the need for splitting the *NFSE-STUR* into two separate instruments is explained. Then, the factor models for each instrument are described and situated in the literature. Once each instrument has been described, the implications and potential impact for each instrument is discussed. Finally, the needs and opportunities for future research are presented.

Using the *NFSE-STUR* to Assess *Understanding and Readiness*

During the development of the survey instrument to measure inservice elementary educators' readiness to implement the new framework for science education, the need to establish both an educator's level of *understanding* about the ideas as well as the *readiness* they felt to implement them became apparent. These separate ideas were assessed by asking teachers to respond to each survey prompt in two ways. First, they

were asked to rate their *understanding* of the idea being presented. Then, they were asked to rate their *readiness* to implement the idea. Committee members agreed that the two were separate constructs and could provide a more complete picture of where teachers are at with the new framework. Running descriptive analyses on the data revealed a noticeable difference between teacher responses in rating their *understanding* versus rating their *readiness* to implement an idea—further strengthening the thought that these measures were assessing different constructs. The exploratory factor analyses conducted on each construct firmly established their distinctness. The responses to *understanding* the ideas basically followed the theme structure identified in the literature and used to develop the items. Explaining the latent factors could be done in a very straightforward manner. However, the factor model for the *readiness* construct emerged with stark differences. Clearly, the only reasonable solution in handling these two powerful and separate constructs is the reorganization of them into two separate instruments. These final instruments are: the *New Framework for Science Education, Survey of Teacher Understanding (NFSE-STU)* and the *New Framework for Science Education, Survey of Teacher Readiness (NFSE-STR)*.

Discussing the Latent Variables for the *Understanding Scale (NFSE-STU)*

Exploratory analysis on the instrument items for the *NFSE-STU* revealed a five-factor model to assess inservice elementary educators' *understanding* of the ideas presented in The Framework. The instrument was initially created based upon six overarching themes that arose from the literature review. These six themes, (1) Science

& Engineering Practices, (2) Crosscutting Concepts, (3) Disciplinary Core Ideas, (4) Integration of the Three Dimensions, (5) Best Practices in Science Education, and (6) Connections to Common Core, are the answer to the first research question. These are the themes developed from the literature that explain the new framework of science education. At the center of this review was *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Research Council, 2012). The first three of these, Science & Engineering Practices, Crosscutting Concepts, and Disciplinary Core Ideas, were lifted straight from the framework as foundational elements,

The committee recommends that science education in grades K-12 be built around three major dimensions. These dimensions are Scientific and engineering practices, Crosscutting concepts that unify the study of science and engineering through their common application across fields, [and] Core Ideas in four disciplinary areas: physical sciences; life sciences; earth and space sciences; and engineering, technology, and applications of science (National Research Council, 2012).

Building from these three, additional overarching ideas were identified as critical to capturing the ideals of The Framework including the integration of the three dimensions, the promotion of best practice in science education, and the emphasis on connecting to the Common Core.

The five-factor model that emerged from the exploratory analysis was nearly identical to the themes used in creating the survey. Table 10 displays the themes and survey items alongside those resulting from the exploratory analysis. Four of the factors were almost exactly the same and the fifth was an easily explained combination of two original themes. Confirmatory factor analysis found that this five-factor model was an acceptable fit to the hypothesized population model.

Table 10. Comparing Understanding Scale Development with the Factor Model

Themes	Instrument Items	PCA Loadings	Factor Name
Science & Engineering Practices	1 – 14	1 – 14 ^a	Science & Engineering Practices
Crosscutting Concepts	15 – 22	15 – 22 ^b	Crosscutting Concepts
Disciplinary Core Ideas	23 – 29	23 – 29, 35 ^c	Teaching Disciplinary Core Ideas
Integration of the Three Dimensions	30 – 34	30 – 34 ^d	Integration of the Three Dimensions
Best Practices in Science Education	35 – 38	36 – 41 ^e	Best Practices in Science Education
Connections to Common Core	39 – 41		

Note. The following items were removed due to multicollinearity or cross-loading:

a. Items 4, 8, and 13

b. Items 15 and 16

c. Item 29

d. Items 30 and 34

e. Items 36 and 41

The first factor loading for the *NFSE-STU* included all of the items in the original Science & Engineering Practices theme; therefore, the title of this factor remained the same. The exploratory analysis revealed that the latent variable is best assessed by eleven of the instrument items. The second factor, Crosscutting Concepts, also loaded identical items as those written to measure the theme. After removing two items, the seven that best assess the factor were kept in the final instrument.

The third factor, Disciplinary Core Ideas, required only a slight adjustment in describing this latent variable. All of the intended items loaded together on this factor;

however, item 35, written for the Best Practices in Science Education theme also loaded with them. Item 35 reads, “When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.” Items 23 – 29 written for this theme, present the NRC’s (2012) vision of the criteria involved in identifying content concepts that are indeed “core” and should be part of a science education. The prompts describe science classes that focus on a few big ideas rather than covering large amounts of content. In addition these prompts further describe science classes as using learning progressions to allow student understanding to develop over time. Curriculum in these classes address big ideas that cut across disciplines—being useful in solving complex problems and relate to life experience. Such classrooms will also be engaging places since understanding core ideas requires experiences with the practices (Michaels et al., 2008). It is no surprise then that item 35, “When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students,” would be aligned with the others. Instead of simply focusing on the core ideas, it appears that the teachers are considering these ideas in terms of their teaching practice which is encouraging as it aligns with foundational literature for The Framework, “Students learn science by actively engaging in the practices of sciences . . .” (National Research Council, 2007, p. 251); therefore, transitioning from theme to latent variable adds just that word, “teaching” to the title, Teaching Disciplinary Core Ideas. Exploratory analysis revealed that seven of the eight items best assessed this variable; those seven were retained.

The three previously discussed themes of The Framework could be seen as the three legs of a stool. An absence of any one of them, and the stool topples. These three ideas are presented by the NRC as interdependent. In developing the NGSS from the framework document, the NGSS Lead States (2013) chose to graphically display this integrated relationship in the logo of the new standards—shown here in Figure 7. The fourth theme used in creating the instrument, then, was the idea that there is an Integration of the Three Dimensions. The exploratory analysis revealed that three of the five items written for this idea were needed to assess the factor.

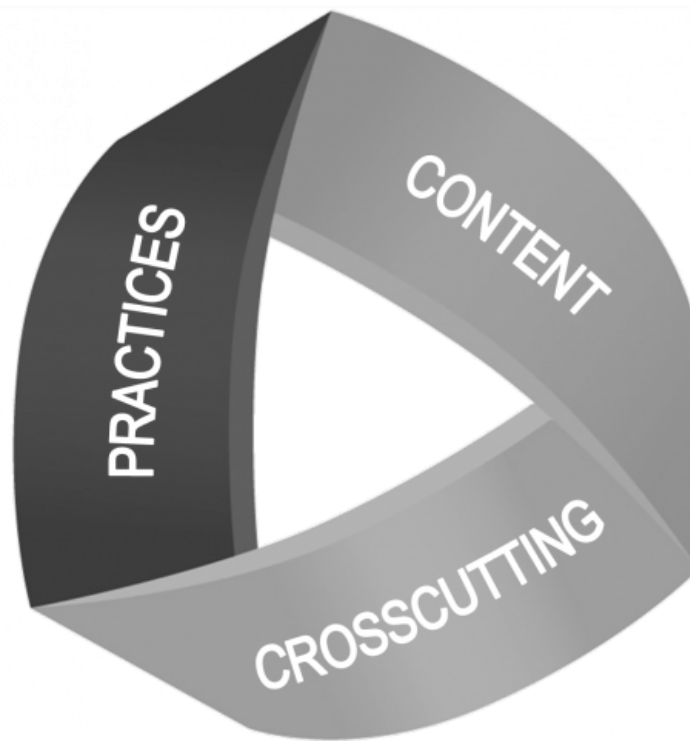


Figure 7. NGSS Logo demonstrating the intended interdependence of the three dimensions.

The fifth and sixth themes used to develop the instrument, Best Practices in Science Education and Connections to Common Core, loaded together during the

exploratory analysis. This was not surprising since the integration of subject areas promoted by Common Core State Standards has long been held to be a best practice as it facilitates the constructivist principles of personal construction of knowledge (Yager & Lutz, 1994). Integration also engenders motivation for learning (D. F. Brown, 2011; MacMath et al., 2010), creates meaningful learning experiences (Beane, 1991; Jacobs, 1989), and can result in higher student achievement (Hartzler, 2000; Romance & Vitale, 2001). The items written for the Common Core connections (39 – 41) describe the need for students to be reading, writing, and speaking about science as well as the importance of mathematical application in scientific investigation and engineering design. Teachers viewed these ideas about subject-area integration as part of the other best practices (a view supported by the literature) presented by items 35 – 38. Therefore, the original title for this factor was kept as originally written, Best Practices in Science Education.

The second research question asked, “What are the underlying dimensions of *NFSE-STUR* items written to assess elementary teachers’ *understanding* of and *readiness* to implement The Framework?” Answering the question necessitated using the instrument items as two separate scales. For the *understanding* scale, the five-factor model that was developed through exploratory analysis was confirmed during the CFA using the 31 items that best described the model. The answer, then, for this question is the valid and reliable final version of the *NFSE-STU* is presented in APPENDIX F.

Discussing the Latent Variables for
the *Readiness Scale (NFSE-STR)*

Exploratory analysis on the instrument items for the *NFSE-STR* revealed a four-factor model to assess inservice elementary educators' *readiness* to implement the ideas presented in The Framework. The instrument was created based upon six themes that arose from the literature review. At the center of this review was *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Research Council, 2012). Unlike the scale measuring *understanding*, results from the factor analysis for the *readiness* scale did not align with the themes identified in the literature. Table 11 displays the initial themes and survey items alongside the factors resulting from exploratory analysis. Only one of the original themes matched a factor explaining how to assess inservice elementary teachers' *readiness* to implement The Framework. Each of the other three factors required redefining. At the same time, the resulting factor model demonstrated higher internal consistencies than the *understanding* scale. The confirmatory factor analysis with the subsample of teacher responses found that, like the *understanding* scale, the fit to the population model was an acceptable fit, yet the *readiness* scale yielded a closer fit than the *understanding* scale.

While the first factor loading for the *NFSE-STR* included most of the items in the original theme, Science & Engineering Practices, there were some key differences. Items 2 and 3 loaded instead with the other items written for the Best Practices in Science Education theme. Plus, items 15 – 17, from the theme, Crosscutting Concepts, loaded with the first factor. Items 15 – 17 address the idea of concepts that cut across disciplines—the concept of patterns and the concept of cause and effect. When looking

at these items together, they can be interpreted as, Students Learning as Scientists & Engineers.

Table 11. Comparing *Readiness Scale* Development with the Factor Model

Themes	Instrument		Factor Name
	Items	PCA Loadings	
Science & Engineering Practices	1 – 14	1, 4 – 17 ^a	Students Learning as Scientists & Engineers
Crosscutting Concepts	15 – 22	18 – 22	Challenging Crosscutting Concepts
Disciplinary Core Ideas	23 – 29	25 – 28,	Integration & Real-World Application of Core Ideas
Integration of the Three Dimensions	30 – 34	30 – 34, 39 – 41 ^b	
Best Practices in Science Education	35 – 38	2 – 3, 23 – 24, 29,	Best Practices for Student Learning
Connections to Common Core	39 – 41	35 – 38 ^c	

Note. The following items were removed due to multicollinearity or cross-loading:

a. Items 4, 13, and 16

b. Items 27 and 31

c. Item 3 and 36

This is the intention of the practices described in the framework. Since scientists and engineers employ the crosscutting concepts in their work, it makes sense that teachers might view these in the same light when thinking about implementing these ideas and having their students learn as scientists and engineers. For example, the most basic of the crosscutting concepts, naturally part of every kindergarten classroom, is “patterns.” Yet, scientists studying the most complex ideas also employ this concept in their practice.

They use patterns to ask scientific questions, determine classifications that can simplify and organize massive numbers of items (National Research Council, 2012).

Why then, did not all of the crosscutting concepts load with the first factor, Students Learning as Scientists & Engineers? A paired samples t test revealed a significant difference between the summed means of items 15 – 17 (representing the crosscutting concepts of patterns and cause and effect) and items 18 – 22 (representing the crosscutting concepts of scale, systems, structure and function, energy and matter, and change). This difference was also practically significant according to the large effect size calculated with Cohen's d . This significant difference demonstrates the change in confidence teachers feel in implementing the ideas. Elementary teachers feel more confidence about implementing the crosscutting concepts presenting in the first three items—cause and effect and patterns. These first two concepts are fundamental in the world of science (National Research Council, 2012). The NRC explains that students begin recognizing patterns prior to coming to school and that cause and effect can be straightforward at times. Evidence from the paired samples t test suggests that teachers also recognize these concepts as more basic. On the other hand, elementary educators feel less confident about implementing the crosscutting concepts of scale, systems, energy and matter, structure and function, and change. Scale is grounded in mathematical literacy, and the final four are interrelated as the first, “systems” is illuminated by the others (National Research Council, 2012). These concepts loading together yet separate from the first concepts can be explained in light of the literature and

could be interpreted as Challenging Crosscutting Concepts, the second factor produced by the exploratory analysis.

The third factor produced by the exploratory analysis best interpreted as Integration & Real-World Application of Core Ideas includes items addressing the Integration of the Three Dimensions theme (Items 30 – 34) and the Connections Common Core theme (Items 39 – 41). The idea threading these items together is integration. They describe unified and meaningful learning experiences that connect the practices of science and engineering with the content of science, the use of concepts that cut across the disciplines of science to unite core ideas, and learning experiences where students are expected to use skills from both mathematics and English language arts.

It makes sense that these items would load together since The Framework embraces integration at a number of levels. It is developed from a recognition that the nature of science is both socially integrated as well as interdisciplinary (National Research Council, 2007). The Framework is also constructed with an appreciation of and need for the integration of knowledge during learning. This idea, a constructivist principle, can be called knowledge assimilation (Piaget & Inhelder, 1969); it is the development of structures in the brain that work in concert to make meaning from new experiences (Rumelhart & Norman, 1981). Case (1991) argued that learning cannot occur without this integration of knowledge. The NRC's (2012) expectation that educators will teach lessons involving all three dimensions: Science & Engineering Practices, Crosscutting Concepts, and Disciplinary Core Ideas embodies this integration of knowledge which occurs where experience and concepts merge (Beijaard & Verloop,

1996). Unlike prior standards documents which separated inquiry as its own standard (National Research Council, 1996), The Framework writers wanted to be sure that changed. They felt strongly enough about this point to include it in their recommendations to future standard's writers. "Recommendation 4: Standards should emphasize all three dimensions articulated in the framework—not only crosscutting concepts and disciplinary core ideas but also science and engineering practices" (National Research Council, 2012, p. 300). The NGSS Lead States (2013) embraced this vision when crafting the NGSS. Each Performance Expectation includes elements from all three dimensions.

Three items focusing on one dimension, the Disciplinary Core Ideas theme, also loaded on this third factor (Items 25 – 28). Why would these load with other items emphasizing the idea of integration? Examining the items in light of the previous discussion indicates the relationship. This third factor revolves around the idea of integration. Items 25 – 28 discuss the disciplinary core ideas as they relate to principles of integration. These items describe a classroom where the focus is on core ideas that build on prior knowledge, have importance across disciplines, and can be applied to real-world situations.

Building on prior knowledge returns to the important idea of integrated knowledge an idea connected strongly to both the literature on integration and The Framework (i.e. Case, 1991; National Research Council, 2012). Core ideas that are important across disciplines promote subject-area integration in a straightforward manner. Core ideas that can be applied to real-world situations address another idea in

integration. In fact connecting school to the real-world is a key aspect of integrated practice. Educators promoting integration historically even described their curriculum as the curriculum of “life” (Whitehead, 1929). A more modern integrationist encourages taking learning into life through social action projects (Kovalik & Olsen, 2005).

Comparing these three items written for disciplinary core ideas alongside the other items loading on this third factor, a relationship emerges and can be explained based on the central idea of integration.

The fourth and final factor resulting from exploratory analysis included items written for the Best Practices in Science Education theme (35 – 38), two items from the Science & Engineering Practices theme (2 – 3), and items addressing the Disciplinary Core Ideas theme (23 – 24, 29). Together, these items highlight a factor best interpreted as, Best Practices for Student Learning. Items loading here describe classrooms where students are actively engaged in their learning, where teachers provide a variety of learning opportunities, and where teachers have their students engage in sustained investigations. Each of these ideas are best practices of science education specifically addressed in The Framework (National Research Council, 2012).

Items 2 and 3, which present the science and engineering practice of questioning, loaded with this factor on best practices since teachers associate questioning with best practice when they consider the application of the idea. This relationship can be explained since questioning is a cognitive practice promoted throughout the disciplines and is not limited to the nature of science. In reading, for example, when readers pose

their own questions, they are engaging in both a cognitive and metacognitive exercise (Perez, 1986). This in turn is a vital part of reading comprehension (Singer, 1978).

Three items addressing disciplinary core ideas (23 – 24, 29) also loaded on this fourth factor. These three items address the idea of focusing on a few core ideas instead of many, the recognition that understanding develops over years not days or weeks, and therefore core ideas to be taught should be investigated at increasing sophistication over multiple years. Each of these ideas align with best practices in education and are not unique to science education. The NRC (1999) addressed each of these in their landmark report on human learning, *How People Learn: Brain, Mind, Experience, and School*.

All of the items loading here on the fourth factor describe ideas about best practice. The items' relationships indicate teachers who are concerned for student learning; therefore, that was added to the title of this factor, Best Practices for Student Learning. Seven of these items were retained since they best assess the latent variable.

The second research question asked, "What are the underlying dimensions of *NFSE-STUR* items written to assess elementary teachers' *understanding* of and *readiness* to implement The Framework?" The answer to that question (for the *readiness* scale) is the 34 item, four-factor *readiness* scale. This scale was established through exploratory analysis and further validated using confirmatory factor analysis. Results from the CFA found that the scale was an acceptable fit to the hypothesized population model and that the four factors yielded high internal consistency. The final version of the *NFSE-STR* is presented in APPENDIX G.

Implications and Applications for Practice

The result of this study is two separate instruments, the *New Framework for Science Education: Survey of Teacher Understanding (NFSE-STU)* and the *New Framework for Science Education: Survey of Teacher Readiness (NFSE-STR)*. The differences in factor models between the two scales present important distinctions between these constructs that align with the literature. The way elementary educators consider their *understanding* about The Framework does not align with the way they describe their *readiness* to implement the ideas. This finding should influence those planning for professional development opportunities. In this section, implications and applications of each and both scales are addressed.

The Uniqueness and Interdependence of the Three Dimensions

The Framework involves a complex interdependence of elements. Each of the three foundational elements, Science & Engineering Practices, Crosscutting Concepts, and Disciplinary Core Ideas, are presented distinctly from each other in the NRC (2012) report. Yet, the need for blending and integrating these three dimensions is also described. Such simultaneous uniqueness and interdependence present special challenges in *understanding*. The exploratory analysis and CFA revealed and confirmed that in-service elementary teachers' recognize the distinctions between these items (the three dimensions in The Framework). This does not mean that teachers understand any of these in depth; but rather, it means that teachers are primed for developing their *understanding* about one or each of these variables. The factor model also demonstrated

that teachers are prepared to develop their *understanding* about the integration of the dimensions; three of the five items, written for the theme, loaded together with high factor loadings of .723, .656, and .585 and a strong internal consistency, $\alpha = .883$. Together these loadings indicate that teachers recognize both the uniqueness and interdependence of these ideas. This model does not describe the depth of inservice teachers' *understanding*, yet it does demonstrate that they are prepared to grow in levels of *understanding* about the complexity these ideas. This also means that teacher *understanding* of these ideas for The Framework can be assessed and professional development planned to meet needs.

However, while teachers are prepared to strengthen their *understanding* about the individuality and integration of each dimension, results from this study indicate that a different approach is needed in preparing teachers to implement these ideas. Results from the factor analysis for items developed for the *readiness* scale, *NFSE-STR*, did not align with the themes identified from the literature. Some crosscutting concepts loaded with the practices, the disciplinary core ideas split between the theme of integration and the theme of best practices, and the theme of the integration of the three dimensions instead organized itself around broader ideas of integration. These differences imply that expecting teachers to implement the individual and integrated elements found in The Framework will require more than cognitive *understanding*.

Connections to Common Core State Standards, Best Practice or Integration?

In the results for the *NFSE-STU*, when describing what they *understand*, teachers view CCSS as part of the best practice latent variable. This is encouraging as it indicates

a conceptual understanding about the role of ELA and mathematics in science education. However, this blended loading also removes the possibility of using the *NFSE-STU* to assess only teacher understandings of CCSS connections in science education.

Adding to the dilemma facing professional developers is the fact that the results from the *NFSE-STR*, the *readiness* for implementation scale, show teachers grouping these ideas differently. Instead of loading with best practices, on the *NFSE-STR*, CCSS loaded in the broad category of integration. However, this distinction is more easily explained or mediated than some other challenges. Since integration can also be described as a best practice, this variation in factor models may be more intellectually interesting than practically important.

Implications for Professional Development

The *NFSE-STU* and *NFSE-STR* were separated during analysis for the very reason that the two explored unique constructs. The factors of the *NFSE-STU* loaded during exploratory analysis nearly as the items were written. This assessed an *understanding* of the elements of The Framework. Interestingly, assessing teachers' *readiness* to implement The Framework produced a factor structure that was almost completely different. These results not provide further evidence that *understanding* and *readiness* are separate constructs; individuals view the relationship of identical ideas differently when employing these two construct "lenses." When describing their *understanding* of the ideas, educators view the connections between items one way—very similar to how the instrument was written. When describing their *readiness* to implement the ideas, their thoughts paint a much different picture. While each of these differences can be

adequately defended, it does leave one planning professional development with questions to be answered. Clearly more research is needed to determine the ramifications of these differences in planning and conducting professional development training for the new framework for science education.

Recommendations for Future Research

This research has developed and validated two survey instruments to measure the current state of readiness elementary teachers self-report in relation to The Framework. While these scales demonstrate a unique effort in the professional development for science education, the research has produced a number of crucial questions needing further research.

At least six direct recommendations for future research can be made based on the study. First, the validation of the *NFSE-STU* and *NFSE-STR* instruments occurred together. Teachers responded to every item in two ways, at the same time. They read the item and responded by rating their level of *understanding* and their level of *readiness*. Nothing about these procedures guarantees that the same results would be achieved when using just one of the final scales in isolation from the other. In fact it is likely that the presence of rating *readiness* artificially inflated participant perception of *understanding*—or vice versa. Validating each scale in isolation should be attempted. Second, data was not collected on participant use of The Framework quotes that were linked to each item. Additional research is needed to determine if the quotes assisted participants and in what way. Third, these instruments need to be validated for middle school and high school science teachers. The current study only validated the instruments for K-5 educators.

Fourth, the new standards, *Next Generation Science Standards: For States, By States* (NGSS Lead States, 2013), were developed from The Framework. It might be hypothesized that the same factor models will form if the instruments were renamed and piloted for NGSS. As of March 2014, eight states have adopted the NGSS. Seeking to validate the instruments in these states could be done in preparation for use in these and other states adopting NGSS in the coming years. Fifth, the instruments could be applied to professional development efforts in training for The Framework. Finally, the results of this study pose a number of critical questions about the similarities and differences between the constructs of understanding and readiness—especially as they apply to professional development in science education. More research is needed before decisions can be made from an empirical position.

Conclusions

The two instruments resulting from this validation research are ready to be piloted and used in like contexts. *NFSE-STU* can be used to assess inservice elementary teachers' *understanding* of the new framework for science education, *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Research Council, 2012). The *NFSE-STR* can be used to assess inservice elementary teachers' *readiness* to implement The Framework in their teaching practice. Though these two instruments were constructed concurrently with identically written items and the same six originating themes, their resulting factor structures produced through exploratory analysis with principal component analysis were vastly different. In spite of small participant numbers, the factor models for both *understanding* and *readiness* scales

were further found to yield acceptable model fit statistics when analyzed using confirmatory factor analysis. These results give the researcher confidence in drawing two conclusions.

First, these instruments can be used to inform professional developers, state agencies, or districts and schools to guide professional development for science education. The *NFSE-STU* will assess levels of teacher *understanding* about The Framework. The results from such assessments will help planners determine areas most appropriate for training to grow levels of *understanding*. If the planners desire to focus instead on *readiness* for implementing The Framework, then the *NFSE-STR*—addressing different underlying constructs—would assist them in their efforts.

A second conclusion that can be addressed in light of this study is that professional developers need to consider their goal and expectations carefully when planning for training on The Framework. This study produced strong evidence, which aligns with elements of the literature (i.e. Beijaard & Verloop, 1996; Parasuraman, 2000), that the constructs of *understanding* and *readiness* are vastly different. Identifying the latent variables for these two scales demonstrated that the differences in these constructs influence the ways in which elementary teachers think about the same set of ideas. When considering their level of *understanding* about particular ideas, teachers' thoughts organized in one certain way. However, when pondering the same ideas in light of their *readiness* to use them in their teaching practice the elements were reorganized into different groups. This means that, training designed to develop teacher understandings of a particular construct could be completely successful in growing teacher understandings

without having a transfer effect in practical application. The results of this study suggest that planners of professional development need to consider carefully the goal of training and plan accordingly.

For example, consider the latent variable, Teaching Disciplinary Core Ideas. This variable was both a theme identified in the literature and used for instrument development as well as a factor that loaded on the *understanding* scale. On the *readiness* scale, the items intended for this theme loaded on two separate factors—each emphasizing a different element—neither centered on the core ideas. In these factors, Integration & Real-World Application of Core Ideas and Best Practices for Student Learning, the core ideas adopted a supporting, blended role in the factor. Failure to consider these complexities in implementing elements of The Framework could greatly limit success in training. If professional developers conducted a training that treated elements of the disciplinary core ideas in isolation from their application in integrated instruction and best practices, teachers' knowledge of the core ideas may grow without any transfer to practical application.

This research takes a step toward a better recognition of elementary teachers' current understandings of The Framework, as well as their current *readiness* to implement its ideals. At the same time, it produced questions to go along with the answers. There is need for further research. In the meantime, professional development planned from either instrument should be done carefully with a focus on understandings about The Framework in the context of practical application.

REFERENCES CITED

- Alreck, P.L., & Settle, R.B. (2004). *The survey research handbook*. New York: McGraw-Hill/Irwin.
- American Association for the Advancement of Science. (1990). *Science for all americans: A project 2061 report on literacy goals in science, mathematics, and technology*. Washinton, DC: Author.
- American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.
- American Educational Research Association. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Anderman, E.M., & Young, A.J. (1994). Motivation and strategy use in science: Individual differences and classroom effects. *Journal of research in science teaching*, 31(8), 811-831. doi: 10.1002/tea.3660310805
- Anderson, J.R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological review*, 94(2), 192.
- Applebee, A.N., Adler, M., & Flihan, S. (2007). Interdisciplinary curricula in middle and high school classrooms: Case studies of approaches to curriculum and instruction. *American Educational Research Journal*, 44(4), 1002-1039.
- Atkin, J.M., & Karplus, R. (1962). Discovery or invention? *The Science Teacher*, 29, 45-51.
- Bandura, A. (1986). *Social functions of thought and action: A social cognitive theory*. Upper Saddle River, NJ: Prentice-Hall.
- Bangert, A.W. (2004). The seven principles of good practice: A framework for evaluating on-line teaching. *The Internet and Higher Education*, 7(3), 217-232. doi: <http://dx.doi.org/10.1016/j.iheduc.2004.06.003>
- Bangert, A.W. (2006). The development of an instrument for assessing online teaching effectiveness. *J. Educational Computing Research*, 35(3), 227-244.
- Bangert, A.W. (2009). Building a validity argument for the community of inquiry survey instrument. *The Internet and Higher Education*, 12, 104-111. doi: 10.1016/j.iheduc.2009.06.001
- Bauer, H.H. (1994). *Scientific literacy and the myth of the scientific method*. Champaign, IL: University of Illinois Press.

- Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental article in science*. Madison, WI: University of Wisconsin Press.
- Beane, J. A. (1991). The middle school: The natural home of integrated curriculum. *Educational Leadership*, 49(2), 9-13.
- Beane, J. A. (1992). Creating an integrative curriculum: Making the connections. *NASSP bulletin*, 76(547), 46-54.
- Beeth, M.E. (1998). Teaching science in fifth grade: Instructional goals that support conceptual change. *Journal of research in science teaching*, 35(10), 1091-1101.
- Beijaard, D., & Verloop, N. (1996). Assessing teachers' practical knowledge. *Studies in educational evaluation*, 22(3), 275-286.
- Bentley, M.L., Ebert, E.S. II, & Ebert, C. (2007). *Teaching constructivist science: Nurturing natural investigations in the standards-based classroom*. Thousand Oaks, CA: SAGE Publications.
- Birenbaum, M. (2007). Evaluating the assessment: Sources of evidence for quality assurance. *Studies in educational evaluation*, 33(1), 29-49.
- Brown, D. F. (2011). Curriculum integration: Meaningful learning based on students' questions. *Middle Grades Research Journal*, 6, 193+.
- Brown, H.E. (1940). The plight of high school physics. *School science and mathematics*, 40(2), 156-160. doi: 10.1111/j.1949-8594.1940.tb04066.x
- Brownell, W. (1933). On the accuracy with which reliability may be measured by correlating test halves. *Journal of Experimental Education*, 1, 204-215.
- Bruce, M.H. (1969). Teacher Education in science. *Review of Educational Research*, 39(4), 415-427. doi: 10.2307/1169706
- Bruner, J.S. (1960). *The process of education*. New York: Vintage.
- Bruner, J.S., Goodnow, J.J., & Austin, G.A. (1956). *A study of thinking*. New York: Wiley.
- Bryne, B. (2009). *Structural equation modeling with AMOS: Basic concepts, applications, and programing* (2nd ed.). London: Routledge.
- Bunting, C. (1987). According to John Dewey: Educational purpose and the new curricula: A view from the theoretical perspective. *NASSP Bulletin*, 71(501), 119-125.

- Bybee, R. (1993). *Reforming science education: Social perspectives and personal reflections*. New York: Teachers College Press.
- Bybee, R. (2002). Scientific inquiry, student learning, and the science curriculum. In R. Bybee (Ed.), *Learning science and the science of learning*. Arlington, VA: NSTA Press.
- Bybee, R. (2006). The National Science Education Standards: Personal reflections. *School science and mathematics, 106*(2), 57-63.
- Bybee, R., Taylor, J.A., Gardner, A., Van Scotter, P., Powell, J.C., Westbrook, A., & Landes, N. (2006). The BSCS 5E instructional model: Origins and effectiveness. A report prepared for the Office of Science Education National Institutes of Health.
- Carey, S. (1991). Knowledge acquisitions: Enrichment or conceptual change? In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 257-291). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Carey, S. (1999). Sources of conceptual change. In E. Scholnick, K. Nelson, S. Gelman & P. Miller (Eds.), *Conceptual Development: Piaget's Legacy* (pp. 293-326). Mahwah, NJ: Lawrence Erlbaum Associates.
- Carey, S., & Smith, C. (1993). On understanding the nature of scientific knowledge. *Educational psychologist, 28*(3), 235-251. doi: 10.1207/s15326985ep2803_4
- Case, R. (1991). The anatomy of curricular integration. *Canadian Journal of Education, 16*(2), 215-224.
- Cattell, J.M. (1914). Science, Education and Democracy. *Science, 39*(996), 154-164. doi: 10.2307/1639778
- Cattell, R.B. (1966). The scree test for the number of factors. *Multivariate behavioral research, 1*(2), 245-276. doi: 10.1207/s15327906mbr0102_10
- Chi, M.T.H. (1992). Conceptual change within and across ontological categories: Examples from learning and discovery in science. In R. Giere (Ed.), *Cognitive models of science: Minnesota studies in the philosophy of science* (pp. 129-186). Minneapolis: University of Minnesota Press.
- Cisterna, Dante, Williams, Michelle, & Merritt, Joi. (2013). Students' understanding of cells & heredity: Patterns of understanding in the context of a curriculum implementation in fifth & seventh grades. *The American Biology Teacher, 75*(3), 178-184. doi: 10.1525/abt.2013.75.3.6

- Clark, L.A., & Watson, D. (1995). Constructing Validity: Basic Issues in Objective Scale Development. *Psychological assessment*, 7(3), 309.
- Cobb, P. (1994). Where Is the mind? constructivist and sociocultural perspectives on mathematical development. *Educational Researcher*, 23(7), 13-20. doi: 10.2307/1176934
- Coleman, D, & Zimba, J. (2008). Math and Science Standards that are Fewer, Clearer, Higher to Raise Achievement at all Levels. A Report prepared for the Carnegie-IAS Commission on Mathematics and Science Education: Retrieved from <http://opportunityequation.org/standards-and-assessments/math-science-standards-are-fewer>.
- Collins, A. (1995). National science education standards in the United States: A process and a product. *Studies in science education*, 26(1), 7-37.
- Collins, A. (2002). How students learn and how teachers teach. In R. Bybee (Ed.), *Learning science and the science of learning* (pp. 3-11). Arlington, VA: NSTA Press.
- Cortina, J.M. (1993). What is coefficient alpha?: An examination of theory and applications. *Journal of applied psychology*, 78(1), 98.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. doi: 10.1007/BF02310555
- Cronbach, L.J., & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological bulletin*, 52(4), 281.
- DeBoer, G.E. (1991). *A history of ideas in science education: Implications for practice*. New York: Teachers College Press.
- Desimone, L.M., & Le Floch, K.C. (2004). Are We Asking the Right Questions? Using Cognitive Interviews to Improve Surveys in Education Research. *Educational Evaluation and Policy Analysis*, 26(1), 1-22. doi: 10.2307/3699501
- DeVellis, R.F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: SAGE Publications.
- Dewey, J. (1902/1966). *The child and the curriculum*. Chicago: The University of Chicago Press.
- Dewey, J. (1938/1988). Experience and Education. In J. Boydston & B. Levine (Eds.), *John Dewey: The Later Works, 1925-1953* (Vol. 13: 1938-1939). Carbondale, IL: Southern Illinois University Press.

- Dewey, J. (1976). The place of manual training in the elementary course of study. In J. Boydston (Ed.), *John Dewey: The Middle Works, 1899-1924* (Vol. 1: 1899-1901, pp. 230-240). Carbondale, IL: Southern Illinois University Press.
- Dilworth, C. (2006). Empiricism Vs. Realism Revisited *The Metaphysics of Science-1* (Vol. 173, pp. 183-192): Springer Netherlands.
- Duschl, R.A. (1990). *Restructuring science education: The importance of theories and their development*. New York: Teachers College Press.
- Dux, P. E., Tombu, M. N., Harrison, S., Rogers, B. P., Tong, F., & Marois, R. (2009). Training improves multitasking performance by increasing the speed of information processing in human prefrontal cortex. *Neuron*, 63(1), 127-138. doi: <http://dx.doi.org/10.1016/j.neuron.2009.06.005>
- Edwards, J.R., & Bagozzi, R.P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological methods*, 5(2), 155.
- Elkind, D. (1967). Editor's introduction. In D. Elkind (Ed.), *Six Psychological Studies*. New York: Random House.
- Ericsson, K.A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *The American psychologist*, 49(8), 725.
- Fabrigar, Leandre R., Wegener, Duane T., MacCallum, Robert C., & Strahan, Erin J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, 4(3), 272-299.
- Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2013). *G*Power* (Version 3.1.6). Germany: Department of Psychology.
- Field, A. (2013). *Discovering statistics using IBM spss statistics*. Thousand Oaks, CA: Sage.
- Fink, A. (2003). *The survey handbook* (2nd ed.). Thousand Oaks, CA: SAGE Publications.
- Fogarty, R. (1991). Ten ways to integrate curriculum. *Educational Leadership*, 49(2), 61.
- Fraley, A. E. (1977). *Core curriculum: An epic in the history of educational reform*. Ann Arbor, MI: University Microfilms International.
- Fulton, J. (1989). *Complete guid to the cognitive process of instruction*. Richmond, VA: Developmental Skills Institute.

- Gehrke, N. J. (1998). A look at curriculum integration from the bridge. *Curriculum Journal*, 9(2), 247-260. doi: 10.1080/0958517970090209
- Gelman, S.A., & Markman, E.M. (1986). Categories and induction in young children. *Cognition*, 23(3), 183-209. doi: [http://dx.doi.org/10.1016/0010-0277\(86\)90034-X](http://dx.doi.org/10.1016/0010-0277(86)90034-X)
- Gewertz, C. (2012). Common standards drive new reading approaches. *Education Week*, 32(12), 2.
- Glaser, R. (1979). Trends and research questions in psychological research on learning nad schooling. *Educational Researcher*, 8, 6-13.
- Göb, Rainer, McCollin, Christopher, & Ramalhoto, MariaFernanda. (2007). Ordinal methodology in the analysis of likert scales. *Quality & Quantity*, 41(5), 601-626. doi: 10.1007/s11135-007-9089-z
- Grant, S. G. (1999). It's just the facts, or is it? An exploration of the relationship between teachers' practices and students' understandings of history.
- Greene, L. C. (1991). Science-centered curriculum in elementary school. *Educational Leadership*, 49(2), 42-46.
- Greenleaf, C. L., Litman, C., Hanson, T. L., Rosen, R., Boscardin, C. K., Herman, J., . . . Jones, B. (2011). Integrating literacy and science in biology: Teaching and learning impacts of reading apprenticeship professional development *American Educational Research Journal*, 48(3), 647-717.
- Hanushek, E.A., Peterson, P.E., & Woessmann, L. (2012). Is the U.S. catching up: international and state trends in student achievement. *Education Next*, 12, 24+.
- Harp, B. (1989). When the principal asks: "How are we using what we know about literacy processes in the content areas?". *The Reading Teacher*, 42(9), 726-727. doi: 10.2307/20200286
- Harris, E.E. (1975). Empiricism in science and philosophy. *Royal Institute of Philosophy supplement*, 9, 154-167. doi: 10.1017/S0080443600001059
- Hartzler, D. S. (2000). *A meta-analysis of studies conducted on integrated curriculum programs and their effects on student achievement*. Indiana University. ProQuest Dissertations and Theses.
- Hawkins, J., & Pea, R.D. (1987). Tools for Bridging the Cultures of Everyday and Scientific Thinking. *Journal of research in science teaching*, 24(4), 291-307.

- Haynes, S.N., Richard, D.C.S., & Kubany, E.S. (1995). Content Validity in Psychological Assessment: A Functional Approach to Concepts and Methods. *Psychological assessment*, 7(3), 238.
- Hodson, D. (2009). *Teaching and learning about science: Language, theories, methods, history, traditions and values*. Boston, MA: Sense Publishers.
- Howe, A.C. (1988). Current issues in science teacher education. *The Clearing House*, 61(7), 309-311. doi: 10.2307/30188350
- Humphrey, G. (1924). The psychology of the Gestalt. *Journal of Educational Psychology*, 15(7), 401.
- Huntley, M. A. (1998). Design and implementation of a framework for defining integrated mathematics and science education. *School Science and Mathematics*, 98(6), 320-327.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. London, England: Routledge and Kagan Paul.
- Jacobs, H. H. (1989). *Interdisciplinary curriculum: Design and implementation*. Alexandria, VA: ASCD.
- Joreskog, K. G., & Sorbom, D. (2001). LISREL (Version 8.51). Chicago: Scientific Software, Inc.
- Khishfe, R., & Abd-El-Khalick, F. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. *Journal of research in science teaching*, 39(7), 551-578.
- Kim, K.H. (2005). The relation among fit indexes, power, and sample size in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 12(3), 368-390.
- Klahr, D. (2000). *Exploring science: The cognition and development of discovery process*. Cambridge, MA: MIT Press.
- Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12(1), 1-48. doi: 10.1207/s15516709cog1201_1
- Klimas, J. (1969). A Cooperative Program for Improving Elementary Science Education. *The American Biology Teacher*, 31(8), 493-496. doi: 10.2307/4442782
- Kovalik, S., & Olsen, K. (2005). Exceeding expectations: A user's guide to implementing brain research in the classroom.

- Krathwohl, D.R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212-218. doi: 10.2307/1477405
- Kuder, G.F., & Richardson, M.W. (1937). The theory of the estimation of test reliability. *Psychometrika, 2*, 151-160.
- Kuhn, D.J. (1973). The science education of the elementary teacher. *School science and mathematics, 73*(6), 463-470. doi: 10.1111/j.1949-8594.1973.tb09081.x
- Kuhn, D.J., & Dean, D. Jr. (2005). Is Developing Scientific Thinking All about Learning to Control Variables? *Psychological Science, 16*(11), 866-870.
- Kuhn, T.S. (1962). *The Structure of scientific revolutions*. Chicago, IL: University of Chicago Press.
- Kulm, G. (2007). Learning from the History of Mathematics and Science Education. *School science and mathematics, 107*(1), 368-368.
- Lawson, A E. (1979). The developmental learning paradigm. *Journal of research in science teaching, 16*(6), 501-515.
- Lederman, N.G. (2011). *The Nature of Science and Inquiry: Their Importance in Developing Scientific Literacy*.
- Lederman, N.G., Abd-El-Khalick, F., Bell, R.L., & Schwartz, R.S. (2002). Views of Nature of Science Questionnaire: Toward Valid and Meaningful Assessment of Learners' Conceptions of Nature of Science. *Journal of research in science teaching, 39*(6), 497-521.
- Ledoux, M., & McHenry, N. (2004). A constructivist approach in the interdisciplinary instruction of science and language arts methods. *Teaching Education, 15*(4), 385-399.
- Leung, W. L. A. (2006). Teaching integrated curriculum: Teachers' challenges. *Pacific Asian Education, 18*(1), 88-102.
- Li, Junlei, Klahr, David, & Siler, Stephanie. (2006). What Lies beneath the Science Achievement Gap: The Challenges of Aligning Science Instruction with Standards and Tests. *Science educator, 15*(1), 1-12.
- Linn, M.C. (1982). Theoretical and practical significance of formal reasoning. *Journal of research in science teaching, 19*(9), 727-742.
- Longino, H. (2002). *The fate of knowledge*. Princeton: Princeton University Press.

- Lonning, R. A., DeFranco, T. C., & Weinland, T. P. (1998). Development of theme-based, interdisciplinary, integrated curriculum: A theoretical model. *School Science and Mathematics, 98*(6), 312-319.
- Lux, N.J. (2010). *Assessing technological pedagogical content knowledge*. (Unpublished doctoral dissertation). Boston University, Boston, MA.
- MacCallum, R.C., Browne, M.W., & Sugawara, H.M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological methods, 1*(2), 130-149.
- MacCallum, R.C., Widaman, K.F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods, 4*(1), 84-99. doi: 10.1037/1082-989X.4.1.84
- MacMath, S., Roberts, J., Wallace, J., & Chi, X. (2010). Curriculum integration and at-risk students: A canadian case study examining student learning and motivation. *British Journal of Special Education, 37*(2), 87-94.
- Mathison, S., & Freeman, M. (1997). *The logic of interdisciplinary studies*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.
- McCoach, D.B., Gable, R.K., & Madura, J.P. (2013). *Instrument development in the affective domain: School and corporate applications* (3rd ed.). Publisher: Springer eBooks.
- McKenna, M. C., & Robinson, R. D. (1990). Content literacy: A definition and implications. *Journal of Reading, 34*(3), 184-186. doi: 10.2307/40014518
- Mestre, J.P., & Cocking, R.R. (2002). Applying the science of learning to the education of prospective science teachers. In R. Bybee (Ed.), *Learning science and the science of learning* (pp. 13-22). Arlington, VA: NSTA Press.
- Metz, K.E. (1995). Reassessment of Developmental Constraints on Children's Science Instruction. *Review of Educational Research, 65*(2), 93-127. doi: 10.2307/1170709
- Metz, S. (2012). New standards. *The Science Teacher, 79*, 6.
- Michaels, S., Shouse, A.W., & Schweingruber, H.A. (2008). *Ready, set, science: Putting research to work in K-8 science classrooms*. Washington, D.C.: National Academies Press.
- Mitchell, E. (1985). The eight-year study: A story not to be forgotten. *The Educational Forum (West Lafayette, Ind.), 50*(1), 11-21.

- Moshman, D. (1982). Exogenous, endogenous, and dialectical constructivism. *Developmental Review*, 2(4), 371-384. doi: [http://dx.doi.org/10.1016/0273-2297\(82\)90019-3](http://dx.doi.org/10.1016/0273-2297(82)90019-3)
- National Aeronautics Space Administration, Greenbelt M. D. Goddard Space Flight Center. (1998). *Understanding our changing planet: NASA's earth science enterprise. 1998 Fact Book*.
- National Center of Educational Statistics. (2013). Trends in International Mathematics and science study (TIMSS). from <http://nces.ed.gov/timss/>
- National Commission on Excellence in Education. (1983). *A nation at risk*. Washington DC: Government Printing Office.
- National Education Association. (1920). *Reorganization of science in secondary schools: A report of the commission on the reorganization of secondary education*. (U.S. Bureau of Education, Bulletin No. 26). Washington, DC: U.S. Government Printing Office.
- National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.
- National Research Council. (1999). *How people learn: Brain, mind, experience, and school*. Washington, D.C.: National Academy Press.
- National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*: National Academy Press.
- National Research Council. (2007). *Taking Science to School: Learning and Teaching Science in Grades K-8*. Washington, DC: The National Academies Press.
- National Research Council. (2012). *A framework for k-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, D.C.: The National Academies Press.
- Navarra, J.G., & Zaffaroni, J. (1960). *Science today for the elementary-school teacher*. New York: Harper & Row.
- Nersessian, N.J. (1989). Conceptual change in science and in science education. *Synthese*, 80(1), 163-183. doi: 10.2307/20116672
- Nersessian, N.J. (2005). Interpreting scientific and engineering practices: Integrating the cognitive, social, and cultural dimensions. In M. Gorman, R. Tweney, D. Gooding & A. Kinconnon (Eds.), *Scientific and technological thinking* (pp. 17-56). Mahwah, NJ: Lawrence Erlbaum Associates.

- Netemeyer, R.G., Bearden, W.O., & Sharma, S. (2003). *Scaling procedures: Issues and applications*. Thousand Oaks, CA: SAGE Publications.
- Newton, D.E. (1968). The dishonesty of inquiry teaching. *School science and mathematics*, 68(9), 807-810. doi: 10.1111/j.1949-8594.1968.tb08364.x
- NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press.
- Nissani, M. (1995). Fruits, salads, and smoothies: A working definition of interdisciplinarity. *Journal of educational thought*, 29(2), 121-128.
- Nixon, R.S., & Barth, K.N. (2014). A comparison of TIMSS items using cognitive domains. *School Science and Mathematics*, 114(2), 65-75. doi: 10.1111/ssm.12054
- Nollmeyer, G.E. (2013). *The integration of English language arts, science and other subjects: Learning from elementary educators' knowledge and practice*. (Unpublished Masters Thesis). Montana State University, Bozeman, MT.
- Norman, D.A. (1980). Twelve issues for cognitive science. *Cognitive Science*, 4(1), 1-32. doi: 10.1207/s15516709cog0401_1
- Nunnally, J.C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Oxford Dictionaries. (Ed.) (2013) Oxford Dictionaries. Oxford University Press.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1), 1-4. doi: 10.1207/S15326985EP3801_1
- Parasuraman, A. (2000). Technology readiness index (Tri): A multiple-item scale to measure readiness to embrace new technologies. *Journal of Service Research*, 2(4), 307-320.
- Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, 116(2), 220-244. doi: 10.1037/0096-1523.19.5.965
- Perez, S.A. (1986). Improving learning through student questioning. *The Clearing House*, 60(2), 63-65. doi: 10.2307/30186620
- Petersen, O.L. (1959). A brief look at the history of science education in America: Its past, present, and future. *Science education (Salem, Mass.)*, 43(5), 427-435.
- Peterson, P. (2013). Revelations from the TIMSS. *Education next*, 13(2).

- Piaget, J. (1967). The mental development of the child. In D. Elkind (Ed.), *Six Psychological Studies*. New York: Random House.
- Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books.
- Pratt, H. (2002). Introduction. In R. Bybee (Ed.), *Learning science and the science of learning* (pp. xiii-xv). Arlington, VA: NSTA Press.
- Provasnik, S., Kastberg, D., Ferraro, D., Lemanski, N., Roey, S., & Jenkins, F. (2012). *Highlights from TIMSS 2011: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context*. Washington, DC: National Center of Educational Statistics, Institute of Educational Sciences, U.S. Department of Education.
- Quinn, H. (2009). What is science? *Physics today*, 62(7), 8.
- Reiter, P.J., & Peinich, J.E. (1991). Elementary science teacher education: Seven exemplars. *Journal of science teacher education*, 2(1), 22-27. doi: 10.1007/BF02614727
- Romance, N. R., & Vitale, M. R. (2001). Implementing an in-depth expanded science model in elementary schools: Multi-year findings, research issues, and policy implications. *International Journal of Science Education*, 23(4), 373-404.
- Romance, N. R., & Vitale, M. R. (2011). *An integrated interdisciplinary model for accelerating student achievement in science and reading comprehension across grades 3-8: Implications for research and practice*. Paper presented at the Society for Research on Educational Effectiveness, Washington D.C.
- Rubba, P.A. (1981). A study of two misconceptions about the nature of science among junior high school students. *School science and mathematics*, 81(3), 221-226.
- Ruchlis, H. (1964). Understanding Elementary Science Education. *School science and mathematics*, 64(5), 388-396.
- Rumelhart, D.E., & Norman, D.A. (1981). *Simulating a skilled typist: A study of skilled cognitive motor performance*.
- Sackes, Mesut. (2010). The role of cognitive, metacognitive, and motivational variables in conceptual change: Preservice early childhood teachers' conceptual understanding of the cause of lunar phases. *ProQuest LLC*.
- Scharmman, LawrenceC, & Orth Hampton, ColleenM. (1995). Cooperative learning and preservice elementary teacher science self-efficacy. *Journal of Science Teacher Education*, 6(3), 125-133. doi: 10.1007/BF02614570

- Schunk, D. (2012). *Learning theories: An educational perspective* (6th ed.). Columbus, OH: Prentice Hall.
- Schwab, J.J. (1962). The teaching of science as enquiry. In J. J. Schwab & P. F. Brandwein (Eds.), *The teaching of science* (pp. 3-103). Cambridge, MA: Harvard University Press.
- Science Education Curriculum, History of. (2010). *Encyclopedia of Curriculum Studies*. Thousand Oaks, CA: SAGE Publications.
- Screen, P. A. (1986). The Warwick Process Science Project. *School Science Review*, 68(242), 12-16.
- Shoemaker, B. J. E. (1991). Education 2000 integrated curriculum. *The Phi Delta Kappan*, 72(10), 793-797.
- Shriner, M., Schlee, B. M., & Libler, R. (2010). Teachers' perceptions, attitudes and beliefs regarding curriculum integration. *Australian Educational Researcher*, 37(1), 51-62.
- Simpson, R.D., & S.J., Oliver. (1990). A summary of major influences on attitude toward and achievement in science among adolescent students. *Science education (Salem, Mass.)*, 74(1), 1-18. doi: 10.1002/sce.3730740102
- Simpson, R.D., & Troost, M.K. (1982). Influences on commitment to and learning of science among adolescent students. *Science education (Salem, Mass.)*, 66(5), 763-781.
- Singer, Harry. (1978). Active comprehension: From answering to asking questions. *The Reading Teacher*, 31(8), 901-908. doi: 10.2307/20194673
- Sirois, S., Spratling, M., Thomas, M. S. C., Westermann, G., Mareschal, D., & Johnson, M. H. (2008). Précis of neuroconstructivism: How the brain constructs cognition. *Behavioral and Brain Sciences*, 31(03), 321-331. doi: doi:10.1017/S0140525X0800407X
- Skinner, B.F. (1950). Are theories of learning necessary? *Psychological review*, 57(4), 193.
- Skinner, B.F. (1975). The steep and thorny way to a science of behavior. *The American psychologist*, 30(1), 42.
- Sneider, C, & Workosky, C. (2009). Critical Issues for Common State Standards Science Anchors Phase I Final Report. A Report Prepared for the National Science Teachers Association: Retrieved from <http://www.nsta.org/pdfs/ScienceAnchorsPhaseOneFinalReport.pdf>.

- Solomon, J., Duveen, J., Scot, L., & McCarthy, S. (1992). Teaching about the nature of science through history: action research in the classroom. *Journal of research in science teaching*, 29(4), 409-421.
- Steele, A., Brew, C., Rees, C., & Ibrahim-Khan, S. (2013). Our practice, their readiness: Teacher educators collaborate to explore and improve preservice teacher readiness for science and math instruction. *Journal of Science Teacher Education*, 24(1), 111-131. doi: 10.1007/s10972-012-9311-2
- Stevens, J.P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Erlbaum.
- SurveyGizmo. (2014). SurveyGizmo Forms & Surveys [Software]. from http://www.surveygizmo.com/?utm_expid=50278599-29.5csRIB26Smq-dZQY9IQ2GQ.0
- Tannenbaum, H.E., Stillman, N., & Piltz, A. (1965). *Science education for elementary school teachers* (2nd ed.). Boston: Allyn and Bacon.
- Tanner, D. (1989). A Brief historical perspective of the struggle for an integrative curriculum. *Educational Horizons*, 68(1), 6-11.
- Thagard, P. (1999). Explanation. In R. Wilson & F. Keil (Eds.), *MIT encyclopedia of cognitive science* (pp. 300-301). Cambridge, MA: MIT Press.
- Thorndike, E.L. (1914). *Educational psychology* (Vol. 2). New York: Teachers College, Columbia University.
- Tienken, C.H. (2013). TIMSS implications for U.S. education. *AASA Journal of Scholarship & Practice*, 9, 3+.
- Tinsley, Howard E., Tinsley, Howard E., & Tinsley, Diane J. (1987). Uses of factor analysis in counseling psychology research. *Journal of counseling psychology*, 34(4), 414-424. doi: 10.1037/0022-0167.34.4.414
- Tippins, D., Nichols, S., & Tobin, K. (1993). Reconstructing science teacher education within communities of learners. *Journal of science teacher education*, 4(3), 65-72. doi: 10.1007/BF02614551
- Toulmin, S. (1982). The construal of reality: Criticism in modern and postmodern science. *Critical Inquiry*, 9(1), 93-111. doi: 10.2307/1343275
- U.S Department of Education. (1997). *Pursuing excellence: A study of U.S. fourth-grade mathematics and science teaching in international context*. Washington DC: National Center for Education Statistics, Government Printing Office.

- U.S. Department of Education. (1995). *Pursuing excellence: A study of U.S. eighth-grade mathematics and science teaching, learning, curriculum, and achievement in international context*. Washington DC: National Center for Education Statistics, Government Printing Office.
- van Merriënboer, J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17(2), 147-177. doi: 10.1007/s10648-005-3951-0
- Vars, G. F. (1991). Integrated curriculum in historical perspective. *Educational Leadership*, 49(2), 14-15.
- Vessel, M.F. (1963). *Elementary school science teaching*. Washington, DC: The Center for Applied Research in Education.
- Victor, E. (1965). *Science for the elementary school*. New York: The MacMillon Company.
- Vitale, M.R., Romance, N.R., & Klentschy, M. (2006). *Improving school reform by changing curriculum policy toward content-area instruction in elementary schools*. Paper presented at the The Annual Meeting of the American Educational Research Association, San Francisco, CA.
- Vosniadou, S., & Brewer, W.F. (1994). Mental models of the day/night cycle. *Cognitive Science*, 18(1), 123-183. doi: 10.1207/s15516709cog1801_4
- Watson, J.B. (1913). Psychology as the behaviorist views it. *Psychological review*, 20, 158.
- Watson, J.B. (1917). The effect of delayed feeding upon learning. *Psychobiology (Baltimore, Md.)*, 1(1), 51.
- Wavering, Michael, Mangione, Katherine, & McBride, Craig. (2013). Reliability of using Piaget's logic of meanings to analyze pre-service teachers' understanding of conceptual problems in earth science: Online Submission.
- Wellington, J. (2001). What is science education for? *Canadian Journal of Science, Mathematics and Technology Education*, 1(1), 23-38. doi: 10.1080/14926150109556449
- West, S.G., Finch, J.F., & Curran, P.J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56-75). Newbury Park, CA: Sage.
- Whitehead, A. N. (1929). *The aims of education and other essays*: Free Press.

- Yager, R.E. (1993). The need for reform in science teacher education. *Journal of science teacher education*, 4(4), 144-148. doi: 10.1007/BF02614568
- Yager, R.E. (2000). The History and Future of Science Education Reform. *The Clearing House*, 74(1), 51-54.
- Yager, R.E. (2005). Toward needed reforms in science teacher education. *Journal of Science Teacher Education*, 16(2), 89-93. doi: 10.1007/s10972-005-2647-0
- Yager, R.E., Lunetta, V.N., & Tamir, P. (1979). Changes in science teacher education 1967-1977. *School science and mathematics*, 79(4), 308-312. doi: 10.1111/j.1949-8594.1979.tb13839.x
- Yager, R.E., & Lutz, M.V. (1994). Integrated science: The importance of 'how' versus 'what'. *School Science and Mathematics*, 94(7), 338.
- Zimmerman, C. (2000). The Development of Scientific Reasoning Skills. *Developmental Review*, 20(1), 99-149. doi: <http://dx.doi.org/10.1006/drev.1999.0497>
- Zimmerman, T.D., & Stage, E.K. (2008). Teaching science for understanding. In L. Darling-Hammond (Ed.), *Powerful Learning: What We Know About Teaching for Understanding* (pp. 151-191). San Francisco, CA: Jossey Bass.

APPENDICES

APPENDIX A

EXPERT REVIEW DRAFT: NFSE-STUR

Construct: Science & Engineering Practices

1. When planning and teaching, educators have students participate in practices used by scientists in the real world.
2. When planning and teaching, educators have students participate in practices used by engineers in the real world.
3. When planning and teaching, educators have students ask questions about scientific phenomena that can drive exploration.
4. When planning and teaching, educators have students ask questions to define engineering problems that can drive design.
5. When planning and teaching, educators have students develop and use models (i.e. mental, conceptual, mathematical, or computer) to explain phenomena.
6. When planning and teaching, educators have students use models to analyze and test existing engineering systems or possible solutions.
7. When planning and teaching, educators have students plan and carry out investigations appropriate to the questions being asked.
8. When planning and teaching, educators have students plan and carry out investigations appropriate to the problem to be solved.
9. When planning and teaching, educators have students organize, analyze, and interpret observational data to make meaning.
10. When planning and teaching, educators have students use mathematical thinking and computational skills to describe and predict phenomena.
11. When planning and teaching, educators have students use mathematical thinking and computational skills to inform engineering design.
12. When planning and teaching, educators have students construct evidence based explanations to describe phenomena that incorporate their understandings about science.
13. When planning and teaching, educators have students design and refine solutions that meet the needs of an engineering problem.
14. When planning and teaching, educators have students participate in discourse and argumentation—using evidence to make their case for either a scientific explanation or an engineering design.
15. When planning and teaching, educators have students obtain and evaluate information from both word and graphic sources.
16. When planning and teaching, educators have students communicate ideas clearly and persuasively through both words and graphics.

Construct: Crosscutting Concepts

17. When planning and teaching, educators use concepts that bridge disciplinary (i.e. physical science, life science, or earth science) boundaries.
18. When planning and teaching, educators have students look for patterns as a strategy in observation or classification.
19. When planning and teaching, educators have students discuss phenomena and develop explanations or designs in terms of cause and effect.
20. When planning and teaching, educators have students develop their understanding of the relationship of scale, proportion, and quantity in the macro and micro world.
21. When planning and teaching, educators have students view portions of the natural world as systems to understand interaction and interdependence.
22. When planning and teaching, educators have students consider the flow and conservation of energy and matter throughout the science curriculum.
23. When planning and teaching, educators have students investigate phenomena in terms of structure and function as a means of sense making.
24. When planning and teaching, educators have students identify things under investigation that remain stable and things that change.

Construct: Disciplinary Core Ideas

25. When planning and teaching, educators focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.
26. When planning and teaching, educators recognize that the development of student understandings of disciplinary core ideas is a progression that takes place over years.
27. When planning and teaching, educators provide opportunities for students to develop an understanding of core ideas in the physical sciences.
28. When planning and teaching, educators provide opportunities for students to develop an understanding of core ideas in the life sciences.
29. When planning and teaching, educators provide opportunities for students to develop an understanding of core ideas in the earth and space sciences.
30. When planning and teaching, educators provide opportunities for students to develop an understanding of core ideas in engineering, technology, and the applications of science.

Construct: Integration of Three Dimensions

31. When planning and teaching, educators weave practices, concepts, and core ideas together into unified experiences.
32. When planning and teaching, educators have students explore disciplinary ideas by engaging in practices and making connections through crosscutting concepts.
33. When planning and teaching, educators intentionally select practices and concepts that best facilitate student sense making for particular core ideas.
34. When planning and teaching, educators have students use the crosscutting concepts when engaging in practices about disciplinary core ideas.
35. When planning and teaching, educators have students apply science and engineering practices within the context of core ideas.

Construct: Best Practices in Science Instruction

36. When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.
37. When planning and teaching, educators develop scientific proficiency by having students participate in a number of scientific activities and thinking.
38. When planning and teaching, educators use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.
39. When planning and teaching, educators have students engage in sustained investigations accompanied by necessary teacher support.

Construct: Connections to Common Core

40. When planning and teaching, educators teach students how to present their scientific ideas and engineering solutions with clarity through both the written and spoken word.
41. When planning and teaching, educators teach students how mathematical concepts and skills apply to scientific investigation and engineering design.
42. When planning and teaching, educators teach students content area reading skills that facilitate their ability to access, interpret, and evaluate scientific knowledge.

APPENDIX B

EXPERT REVIEW STAGE TWO POWER POINT

Thanks for Joining Us

This will be very informal
It won't take more than an hour

Fundamental Issue with framework quotes

Quote from reviewer:

I think the problem with many of the Framework quotes is that they already assume sophisticated thinking about teaching science and the nature of science. The quotes are "clear" to someone that is already thinking in this direction...but not very helpful for someone new to these ideas.

Separating Science & Engineering Practices

Where separation is questioned:

- Practice of Planning & Carrying Out Investigations
- Practice of Mathematical Thinking

Models

The 2 Models Prompts:

- When planning and teaching, educators have students develop and use models (i.e. mental, conceptual, mathematical, or computer) to explain phenomena.
- When planning and teaching, educators have students use models to analyze and test existing engineering systems or possible solutions.

Suggestions from you:

- Multiple questions for different models
- Change quote to describe types
- Change prompts/quote to focus on mental and conceptual models
- Remove engineering prompt

Practice of Mathematical Thinking

The 2 Math Prompts:

- When planning and teaching, educators have students use mathematical thinking and computational skills to describe and predict phenomena.
- When planning and teaching, educators have students use mathematical thinking and computational skills to inform engineering design

Generally low numbers on clarity & conciseness

- We already addressed combining science and engineering
- Separate mathematical thinking and computational thinking

Issues of Scale

The Prompt:

- When planning and teaching, educators have students develop their understanding of the relationship of scale, proportion, and quantity in the macro and micro world.

Rewritten based on suggestions:

- . . . have students develop an understanding of scale, proportion, and quantity
- . . . have students develop their understanding that phenomena work differently at different scales

Disciplinary Core Ideas

The Framework proposes core ideas for the purpose of unifying content around conceptual ideas that can develop through progressions over time. And so that the sheer amount of content is limited.

How do I write prompts to get at these ideas?

Current DCI Prompts

- When planning and teaching, educators focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.
- When planning and teaching, educators recognize that the development of student understandings of disciplinary core ideas is a progression that takes place over years.
- When planning and teaching, educators provide opportunities for students to develop an understanding of core ideas in the physical sciences.
- When planning and teaching, educators provide opportunities for students to develop an understanding of core ideas in the life sciences.
- When planning and teaching, educators provide opportunities for students to develop an understanding of core ideas in the earth and space sciences.
- When planning and teaching, educators provide opportunities for students to develop an understanding of core ideas in engineering, technology, and the applications of science.

Best Practice

Eliminate this Section or Rework it?
The prompts:

- When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.
- When planning and teaching, educators develop scientific proficiency by having students participate in a number of scientific activities and thinking.
- When planning and teaching, educators use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.
- When planning and teaching, educators have students engage in sustained investigations accompanied by necessary teacher support.

Prompts Needing Comment

These prompts received low score on clarity and/or conciseness but no specific comments were attached:

- When planning and teaching, educators have students design and refine solutions that meet the needs of an engineering problem.
- When planning and teaching, educators have students participate in discourse and argumentation—using evidence to make their case for either a scientific explanation or an engineering design.

APPENDIX C

PILOT DRAFT: NFSE-STUR

Construct: Science & Engineering Practices

1. When planning and teaching, educators have students participate in practices used by scientists and engineers in the real world.
2. When planning and teaching, educators have students ask questions about scientific phenomena that can drive exploration.
3. When planning and teaching, educators have students ask questions to define engineering problems that can drive design.
4. When planning and teaching, educators have students develop and refine mental models as part of the sense making process.
5. When planning and teaching, educators have students develop and refine conceptual models to express their understanding about scientific phenomena.
6. When planning and teaching, educators have students develop models to visualize and refine an engineered design
7. When planning and teaching, educators have students plan and carry out investigations to gather data about scientific phenomena and engineering problems.
8. When planning and teaching, educators have students organize, analyze, and interpret observational data to make meaning.
9. When planning and teaching, educators have students apply mathematical and computational thinking to investigate scientific questions and engineering problems.
10. When planning and teaching, educators have students construct evidence-based explanations to describe phenomena that incorporate their understandings about science.
11. When planning and teaching, educators have students design and refine solutions that meet the needs of an engineering problem.
12. When planning and teaching, educators have students engage in evidence-based argumentation about scientific explanations and engineered designs.
13. When planning and teaching, educators have students obtain and evaluate information from words, images, and other media.
14. When planning and teaching educators have students communicate ideas clearly and persuasively through words, images, and other media.

Construct: Crosscutting Concepts

15. When planning and teaching, educators use concepts that bridge disciplinary (i.e. physical science, life science, or earth science) boundaries.

16. When planning and teaching, educators have students look for patterns as a strategy in observation or classification.
17. When planning and teaching, educators have students consider issues of cause and effect when questioning and discussing scientific phenomena or engineering designs.
18. When planning and teaching, educators have students develop an understanding that phenomena work differently at different scales.
19. When planning and teaching, educators have students use systems thinking when investigating scientific phenomena.
20. When planning and teaching, educators have students consider that since energy and matter are conserved, much can be determined by studying their flow into and out of systems.
21. When planning and teaching, educators have students investigate phenomena in terms of structure and function as a means of sense making.
22. When planning and teaching, educators have students identify what aspects of a system remain stable over time and what aspects undergo patterns of change.

Construct: Disciplinary Core Ideas

23. When planning and teaching, educators focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.
24. When planning and teaching, educators recognize that the development of student understandings of disciplinary core ideas is a progression that takes place over years.
25. When planning and teaching, educators use a curriculum that follows a progression so that students engage in learning experiences that build on what they have previously learned and that prepare them for what they need to learn later.
26. When planning and teaching, educators include core ideas that have broad importance across multiple disciplines or are key organizing principles within a discipline.
27. When planning and teaching, educators include core ideas that are important in investigating more complex ideas and solving problems.
28. When planning and teaching, educators include core ideas that relate to the interests and life experiences of students or societal concerns.
29. When planning and teaching, educators include core ideas that can be investigated at multiple grade levels with increasing sophistication.

Construct: Integration of Three Dimensions

30. When planning and teaching, educators weave practices, concepts, and core ideas together into unified experiences.
31. When planning and teaching, educators have students explore disciplinary ideas by engaging in practices and making connections through crosscutting concepts.
32. When planning and teaching, educators intentionally select practices and concepts that best facilitate student sense making for particular core ideas.
33. When planning and teaching, educators have students use the crosscutting concepts when engaging in practices about disciplinary core ideas.
34. When planning and teaching, educators have students apply science and engineering practices within the context of core ideas.

Construct: Best Practices in Science Instruction

35. When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.
36. When planning and teaching, educators develop scientific proficiency by having students participate in a number of scientific activities and thinking.
37. When planning and teaching, educators use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.
38. When planning and teaching, educators have students engage in sustained investigations accompanied by necessary teacher support.

Construct: Connections to Common Core

39. When planning and teaching, educators teach students how to present their scientific ideas and engineering solutions with clarity through both the written and spoken word.
40. When planning and teaching, educators teach students how mathematical concepts and skills apply to scientific investigation and engineering design.
41. When planning and teaching, educators teach students content area reading skills that facilitate their ability to access, interpret, and evaluate scientific knowledge.

APPENDIX D

VALIDATION DRAFT: NFSE-STUR

Construct: Science & Engineering Practices

1. When planning and teaching, educators have students participate in practices used by scientists and engineers in the real world.
2. When planning and teaching, educators have students ask questions about scientific phenomena that can drive exploration.
3. When planning and teaching, educators have students ask questions to define engineering problems that can drive design.
4. When planning and teaching, educators have students develop and refine mental models as part of the sense making process.
5. When planning and teaching, educators have students develop and refine conceptual models to express their understanding about scientific phenomena.
6. When planning and teaching, educators have students develop models to visualize and refine an engineered design
7. When planning and teaching, educators have students plan and carry out investigations to gather data about scientific phenomena and engineering problems.
8. When planning and teaching, educators have students organize, analyze, and interpret observational data to make meaning.
9. When planning and teaching, educators have students apply mathematical and computational thinking to investigate scientific questions and engineering problems.
10. When planning and teaching, educators have students construct evidence-based explanations to describe phenomena that incorporate their understandings about science.
11. When planning and teaching, educators have students design and refine solutions that meet the needs of an engineering problem.
12. When planning and teaching, educators have students engage in evidence-based argumentation about scientific explanations and engineered designs.
13. When planning and teaching, educators have students obtain and evaluate information from words, images, and other media.
14. When planning and teaching educators have students communicate ideas clearly and persuasively through words, images, and other media.

Construct: Crosscutting Concepts

15. When planning and teaching, educators use concepts that bridge disciplinary (i.e. physical science, life science, or earth science) boundaries.

16. When planning and teaching, educators have students look for patterns as a strategy in observation or classification.
17. When planning and teaching, educators have students consider issues of cause and effect when questioning and discussing scientific phenomena or engineering designs.
18. When planning and teaching, educators have students develop an understanding that phenomena work differently at different scales.
19. When planning and teaching, educators have students use systems thinking when investigating scientific phenomena.
20. When planning and teaching, educators have students consider that since energy and matter are conserved, much can be determined by studying their flow into and out of systems.
21. When planning and teaching, educators have students investigate phenomena in terms of structure and function as a means of sense making.
22. When planning and teaching, educators have students identify what aspects of a system remain stable over time and what aspects undergo patterns of change.

Construct: Disciplinary Core Ideas

23. When planning and teaching, educators focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.
24. When planning and teaching, educators recognize that the development of student understandings of disciplinary core ideas is a progression that takes place over years.
25. When planning and teaching, educators use a learning progression approach by building from prior knowledge and working towards future sophistication.
26. When planning and teaching, educators include core ideas that have broad importance across multiple disciplines or are key organizing principles within a discipline.
27. When planning and teaching, educators include core ideas that are important in investigating more complex ideas and solving problems.
28. When planning and teaching, educators include core ideas that relate to the interests and life experiences of students or societal concerns.
29. When planning and teaching, educators include core ideas that can be investigated at multiple grade levels with increasing sophistication.

Construct: Integration of Three Dimensions

30. When planning and teaching, educators weave practices, concepts, and core ideas together into unified experiences.
31. When planning and teaching, educators have students explore disciplinary ideas by engaging in practices and making connections through crosscutting concepts.
32. When planning and teaching, educators intentionally select practices and concepts that best facilitate student sense making for particular core ideas.
33. When planning and teaching, educators have students use the crosscutting concepts when engaging in practices about disciplinary core ideas.
34. When planning and teaching, educators have students apply science and engineering practices within the context of core ideas.

Construct: Best Practices in Science Instruction

35. When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.
36. When planning and teaching, educators develop scientific proficiency by having students participate in a number of scientific activities and thinking.
37. When planning and teaching, educators use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.
38. When planning and teaching, educators have students engage in sustained investigations accompanied by necessary teacher support.

Construct: Connections to Common Core

39. When planning and teaching, educators teach students how to present their scientific ideas and engineering solutions with clarity through both the written and spoken word.
40. When planning and teaching, educators teach students how mathematical concepts and skills apply to scientific investigation and engineering design.
41. When planning and teaching, educators teach students content area reading skills that facilitate their ability to access, interpret, and evaluate scientific knowledge.

APPENDIX E

EXPERT REVIEW STAGE ONE COMMENTS

On the memberships, might you put in examples (NSTA, NCTM, etc)? What about a question like, What is your school's/district's commitment to science education at your grade level? Then offer choices like no commitment/little commitment/a degree of commitment/fully committed...or something like that. ??? How detailed do you think the final question will be? Again, should there be examples?

Could you include familiarity with NGSS or other NRC publications? Perhaps that is later? (Maybe specifically ask if they have read Ready Set Science or America's Lab Report?) The other demographics look fine.

Are you limiting your participants to only include current science teachers? If not, some of the questions don't make sense. I assume you will be giving them an actual scale to use in all the places you ask them to rate things. I would give them some options to check off which organizations they belong to, along with the option of including others. I would give them some options to check off for the in-service PD they have received, along with the option of including others.

- Please rate your current enjoyment in teaching science.

- Please rate how successful you feel in teaching science.

I would have had an easier time answering these questions with a scale.

- Please rate your familiarity with current thoughts about science teaching (i.e. teaching science inquiry, incorporating the science and engineering practices, or using learning progressions).

If you want to get specific answers from each part you may want to break it down into multiple questions

1. Should this be YOU and not educators?
By "educators" are you asking if the person taking the survey does this...or educators in general? This same comment probably fits for other questions... I would edit to say "... educators should have students engage in practices..." (The addition of "should" is something that I think should be universally applied to all the statements.) I don't know how important the phrase "When planning and teaching" is. I would think it could be removed. (This applies to all the statements.)
2. If you change the prompt to YOU above, it should be changed everywhere. This makes it more personal. You should think about whether or not you want what teachers ARE doing or what they SHOULD be doing. I'm thinking you want to know they are doing. If you want both, then you could do the following: 2. When planning and teaching, YOU have students participate in practices used by engineers in the real world. 2. My belief is that when planning and teaching, educators have students participate in practices used by engineers in the real world. or reverse the order...it depends on your goal for the questions.

It took me a couple of times reading this question to finally pick up that this was related to engineers. Maybe underline that word.

I don't think this item needs to be separate from the one above it. I would make just one statement with both scientists and engineers included.

- 3.
4. Same as question 2. Highlight that this is engineering. I'm sure they will have it figured out by now.
Change to "Educators should have students identify and define problems that can drive the design process."
5. the use of models is very open ended -- I think multiple questions might be needed here -- to identify different types of models.
Models were initially confusing to me as I worked through the framework. It is probably good to include different examples of models. The quote doesn't contain types of models and so if you include it in the question you may need to include it here as well.
From page 56-57 of the Framework: Scientists construct mental and conceptual models of phenomena. Mental models are internal, personal, idiosyncratic, incomplete, unstable, and essentially functional. They serve the purpose of being a tool for thinking with, making predictions, and making sense of experience. Conceptual models, the focus of this section, are, in contrast, explicit representations that are in some ways analogous to the phenomena they represent. ... Conceptual models are in some senses the external articulation of the mental models that scientists hold and are strongly interrelated with mental models. Building an understanding of models and their role in science helps students to construct and revise mental models of phenomena. Better mental models, in turn, lead to a deeper understanding of science and enhanced scientific reasoning. Scientists use models (from here on, for the sake of simplicity, we use the term "models" to refer to conceptual models rather than mental models) to represent their current understanding of a system (or parts of a system) under study, to aid in the development of questions and explanations, and to communicate ideas to others.
6. same as before -- models can be interpreted in many ways.
This use of modeling for an engineering purpose is sufficiently different from its use in science, that I can see having it as a different prompt. That said, the necessity of a teacher understanding this is not as great as it is for the idea of modeling in a science context.
7. It may also be helpful to use the NGSS grade band expectations for the practices -- throughout.

You could include the word science (or scientific investigation) in this statement.

I think that it would be nice if the prompt and/or the NRC quote emphasized the idea that the purpose of investigations is to collect data (observations, measurements, etc.) which can be used as evidence to answer a question. I am not sure if the phrase "appropriate to" is as clear and concise as it could be.

- 8 You may want to include engineering problem in this statement.
I would combine this with the one above. The use of investigations in science and in engineering are not sufficiently different to have two separate prompts.
- 9 This is an example where the framework quote focuses on grade 12 end point ...that is not appropriate for elementary. NGSS endpoints might be better. I'm not sure that adding "organize" to the list adds anything, but it probably doesn't hurt either.
- 10 again the NGSS endpoints may provide more clarity.
I think the phrase "mathematical thinking and computational skills" is misleading. I would go back to the phrase in the Framework and NGSS of "use mathematics and computational thinking" In fact, it might be worthwhile to separate mathematics and computational thinking into two separate prompts, as I would bet that while most teachers know what mathematics is, many would not feel comfortable with defining "computational thinking"
This is a mystery to me on how to do with certain science phenomenon
11. NGSS endpoints...
There is no reason to have this separate from the question on mathematical thinking for science.
- 12 An elementary teacher will really struggle with this quote.
- 13 From page 68-69 of the Framework: In engineering, the goal is a design rather than an explanation. The process of developing a design is iterative and systematic, as is the process of developing an explanation or a theory in science. Engineers'™ activities, however, have elements that are distinct from those of scientists. These elements include specifying constraints and criteria for desired qualities of the solution, developing a design plan, producing and testing models or prototypes, selecting among alternative design features to optimize the achievement of design criteria, and refining design ideas based on the performance of a prototype or simulation.
- 14 I wish the quote from the NRC included the idea of defending models in addition to defending explanations.

- 15 word sources sounds awkward. Perhaps scientific texts would suffice. The phrase "from both word and graphic sources." is very awkward. I suggest replacing it with "from verbal, visual, and multimedia sources."
- 16 I think the question is one that educators will be comfortable with, so the quote doesn't add much- words and graphics do not contain the oral distribution of knowledge. The phrase "through both words and graphics." is very awkward. I suggest replacing it with "using words, images and other media."
- 17 My comments about wording are the same for this section. That's the reason for the 2 in the clarity category. This will be a hard question if they do not have exposure to the concept of ccc. Should this include engineering as well? I don't think that this point is very important. If included, I would emphasize the point of applying to many different disciplines rather than using the term bridging. (and yes, I realize the Framework uses that term, but I don't think it is very clear.)
- 18 The NGSS endpoints may be more helpful for all of these questions... Patterns may be confusing to some educators. It is not defined well in the statement or the quote.
- 19 Suggested revision: Educators should have students consider issues of cause and effect when they discuss phenomena and develop explanations. Educators should have students frame their descriptions and explanations of phenomena in terms of cause and effect.
- 20 Micro doesn't cover everything outside of the macroscopic. I may be confused on this point. Educators should have students develop their understanding of scale, proportion, and quantity. Educators should have students develop their understanding that phenomena work differently at different scales. From page 89 of the Framework: The understanding of relative magnitude is only a starting point. As noted in Benchmarks for Science Literacy, "The large idea is that the way in which things work may change with scale. Different aspects of nature change at different rates with changes in scale, and so the relationships among them change, too" [4]. Appropriate understanding of scale relationships is critical as well to engineering "no structure could be conceived, much less constructed, without the engineer's precise sense of scale.
- 21 Examples might be helpful.

Educators should have students use systems thinking when analyzing and explaining phenomena. From pages 91-92 of the Framework As noted in the National Science Education Standards, "The natural and designed world is complex; it is too large and complicated to investigate and comprehend all at once. Scientists and students learn to define small portions for the convenience of investigation. The units of investigations can be referred to as "systems." A system is an organized group of related objects or components that form a whole. Systems can consist, for example, of organisms, machines, fundamental particles, galaxies, ideas, and numbers. Systems have boundaries, components, resources, flow, and feedback •

- 22 Educators should have students consider that since energy and matter are conserved, much can be determined by studying their flow into and out of systems.
- 23 Educators should have students consider structure and function in developing descriptions and explanations of phenomena. Typo in the Framework quote: "build" should be "built"
- 24 I think the problem with many of the Framework quotes is that they already assume sophisticated thinking about teaching science and the nature of science. The quotes are "clear" to someone that is already thinking in this direction...but not very helpful for someone new to these ideas.
Educators should have students identify what aspects of a system remain stable over time and what aspects undergo patterns of change. The NRC Quote describes stability, but not patterns of change. This is a very long section in the Framework and I think you are going to have to carefully pull several sentences from throughout to cover all the necessary ground. BTW: I think it would be better to call these "Framework Quotes" rather than "NRC Quote"
Educators should know about the Framework, but don't necessarily need to know it comes from the NRC.
- 25 Same rationale here...
This is ambiguous.... but I don't know how to avoid that. What does it mean to focus on a few core ideas vs. large number? ...it depends!
I would change the word "topics" to the word "facts"
- 26 Is it possible to do some interpretation here? Perhaps focus on a teacher's understanding of what comes before and what comes next in terms of developing these ideas. It is easy to answer this one "yes" without actually putting it into practice.
Educators should use a curriculum that follows a progression so that students engage in learning experiences that build on what they have learned earlier and that prepare them for what they need to learn later.

- 27 providing opportunities does not mean that it is sufficient or effective...
The NRC quote may contain a little too much specific information. I'm sure that this is a delicate balancing act.
I don't think there is much to be gained asking whether its important to learn each discipline.
- 28 providing opportunities does not mean that it is sufficient or effective...
The NRC quote may contain a little too much specific information. I'm sure that this is a delicate balancing act.
I don't think there is much to be gained asking whether its important to learn each discipline.
- 29 providing opportunities does not mean that it is sufficient or effective...
The NRC quote may contain a little too much specific information. I'm sure that this is a delicate balancing act.
I don't think there is much to be gained asking whether its important to learn each discipline.
- 30 providing opportunities does not mean that it is sufficient or effective...
The NRC quote may contain a little too much specific information. I'm sure that this is a delicate balancing act.
I don't think there is much to be gained asking whether its important to learn each discipline.
- 31 Same rationale
Question 32 covers all the ground that is needed for this section.
- 32 I am having a hard time differentiating between 31 and 32. This could be me.
Question 32 covers all the ground that is needed for this section.
- 33 This is a tough question -- I like it, but wonder if a second could be added about developing practices (explicitly) throughout the year.
Student sense making is unclear and it is not covered in the NRC quote.
Question 32 covers all the ground that is needed for this section.
- 34 again -- something about developing these connections throughtout the year.
Question 32 covers all the ground that is needed for this section.
- 35 PErhaps this question can be modified to focus on developing practices throughout the school year or course.
Question 32 covers all the ground that is needed for this section.
- 36 Same scoring rationale

I don't know how helpful this question is -- Although teachers may not "do" this, I can't imagine that they would answer no.

Is active participation the same as actively involved in the kinds of learning that

- 37 This is another one of those questions that may not be helpful. I can imagine most people will answer yes / strongly agree -- even if they don't do it...or don't know how to do it.
The word thinking confuses me here.
This seems redundant with the sections about the practices.
- 38 To me, the question and quote could just be written as educators do stuff.... It is a throw-away question.
student-led rather than students-led
This doesn't seem particularly noteworthy.
- 39 What does it mean to be sustained? 30 minutes? 2 days? weeks? etc.
This doesn't seem particularly noteworthy.
- 40 Same scoring rationale here
This seems like it should be with the practices / practice 8 / questions.
Sometimes difficult to assess, but necessary for students to do
- 41 This seems repetitive from the practices questions
These items are redundant with the section on the practices.
- 42 repeats questions about practice 8.
These items are redundant with the section on the practices.

APPENDIX F

NEW FRAMEWORK FOR SCIENCE EDUCATION:
SURVEY OF TEACHER UNDERSTANDING

New Framework for Science Education: Survey of Teacher Understanding

Science & Engineering Practices

1. When planning and teaching, educators have students participate in practices used by scientists and engineers in the real world.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

2. When planning and teaching, educators have students ask questions about scientific phenomena that can drive exploration.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

3. When planning and teaching, educators have students ask questions to define engineering problems that can drive design.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

4. When planning and teaching, educators have students develop and refine conceptual models to express their understanding about scientific phenomena.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

5. When planning and teaching, educators have students develop models to visualize and refine an engineered design.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

6. When planning and teaching, educators have students plan and carry out investigations to gather data about scientific phenomena and engineering problems.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

7. When planning and teaching, educators have students apply mathematical and computational thinking to investigate scientific questions and engineering problems.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

8. When planning and teaching, educators have students construct evidence-based explanations to describe phenomena that incorporate their understandings about science.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

9. When planning and teaching, educators have students design and refine solutions that meet the needs of an engineering problem.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

10. When planning and teaching, educators have students engage in evidence-based argumentation about scientific explanations and engineered designs.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

11. When planning and teaching educators have students communicate ideas clearly and persuasively through words, images, and other media.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

Teaching Disciplinary Core Ideas

12. When planning and teaching, educators focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

13. When planning and teaching, educators recognize that the development of student understandings of disciplinary core ideas is a progression that takes place over years.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

14. When planning and teaching, educators use a learning progression approach by building from prior knowledge and working towards future sophistication.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

15. When planning and teaching, educators include core ideas that have broad importance across multiple disciplines or are key organizing principles within a discipline.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

16. When planning and teaching, educators include core ideas that are important in investigating more complex ideas and solving problems.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

17. When planning and teaching, educators include core ideas that relate to the interests and life experiences of students or societal concerns.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

18. When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------



Crosscutting Concepts

19. When planning and teaching, educators have students consider issues of cause and effect when questioning and discussing scientific phenomena or engineering designs.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	-------------------------	--------------------	------------------------	-------------------------	---------------------------

20. When planning and teaching, educators have students develop an understanding that phenomena work differently at different scales.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	-------------------------	--------------------	------------------------	-------------------------	---------------------------

21. When planning and teaching, educators have students use systems thinking when investigating scientific phenomena.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	-------------------------	--------------------	------------------------	-------------------------	---------------------------

22. When planning and teaching, educators have students consider that since energy and matter are conserved, much can be determined by studying their flow into and out of systems.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	-------------------------	--------------------	------------------------	-------------------------	---------------------------

23. When planning and teaching, educators have students investigate phenomena in terms of structure and function as a means of sense making.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	-------------------------	--------------------	------------------------	-------------------------	---------------------------

24. When planning and teaching, educators have students identify what aspects of a system remain stable over time and what aspects undergo patterns of change.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

Integration of the Three Dimensions

25. When planning and teaching, educators have students explore disciplinary ideas by engaging in practices and making connections through crosscutting concepts.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

26. When planning and teaching, educators intentionally select practices and concepts that best facilitate student sense making for particular core ideas.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

27. When planning and teaching, educators have students use the crosscutting concepts when engaging in practices about disciplinary core ideas.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

Best Practices in Science Education

28. When planning and teaching, educators use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

29. When planning and teaching, educators have students engage in sustained investigations accompanied by necessary teacher support.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

30. When planning and teaching, educators teach students how to present their scientific ideas and engineering solutions with clarity through both the written and spoken word.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

31. When planning and teaching, educators teach students how mathematical concepts and skills apply to scientific investigation and engineering design.

No Understanding	Slight Understanding	Fair Understanding	Solid Understanding	Strong Understanding	Advanced Understanding
------------------	----------------------	--------------------	---------------------	----------------------	------------------------

APPENDIX G

NEW FRAMEWORK FOR SCIENCE EDUCATION:
SURVEY OF TEACHER READINESS

Students Learning as Scientists & Engineers

1. When planning and teaching, educators have students participate in practices used by scientists and engineers in the real world.

No Readiness Slight Readiness Fair Readiness Solid Readiness Strong Readiness Advanced Readiness

2. When planning and teaching, educators have students develop and refine conceptual models to express their understanding about scientific phenomena.

No Readiness Slight Readiness Fair Readiness Solid Readiness Strong Readiness Advanced Readiness

3. When planning and teaching, educators have students develop models to visualize and refine an engineered design.

No Readiness Slight Readiness Fair Readiness Solid Readiness Strong Readiness Advanced Readiness

4. When planning and teaching, educators have students plan and carry out investigations to gather data about scientific phenomena and engineering problems.

No Readiness Slight Readiness Fair Readiness Solid Readiness Strong Readiness Advanced Readiness

5. When planning and teaching, educators have students organize, analyze, and interpret observational data to make meaning.

No Readiness Slight Readiness Fair Readiness Solid Readiness Strong Readiness Advanced Readiness

6. When planning and teaching, educators have students apply mathematical and computational thinking to investigate scientific questions and engineering problems.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

7. When planning and teaching, educators have students construct evidence-based explanations to describe phenomena that incorporate their understandings about science.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

8. When planning and teaching, educators have students design and refine solutions that meet the needs of an engineering problem.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

9. When planning and teaching, educators have students engage in evidence-based argumentation about scientific explanations and engineered designs.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

10. When planning and teaching educators have students communicate ideas clearly and persuasively through words, images, and other media.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

11. When planning and teaching, educators use concepts that bridge disciplinary (i.e. physical science, life science, or earth science) boundaries.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

12. When planning and teaching, educators have students consider issues of cause and effect when questioning and discussing scientific phenomena or engineering designs.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

Integration & Real-World Application of Core Ideas

13. When planning and teaching, educators use a learning progression approach by building from prior knowledge and working towards future sophistication.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

14. When planning and teaching, educators include core ideas that have broad importance across multiple disciplines or are key organizing principles within a discipline.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

15. When planning and teaching, educators include core ideas that relate to the interests and life experiences of students or societal concerns.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

16. When planning and teaching, educators weave practices, concepts, and core ideas together into unified experiences.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

17. When planning and teaching, educators intentionally select practices and concepts that best facilitate student sense making for particular core ideas.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

18. When planning and teaching, educators have students use the crosscutting concepts when engaging in practices about disciplinary core ideas.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

19. When planning and teaching, educators have students apply science and engineering practices within the context of core ideas.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

20. When planning and teaching, educators teach students how to present their scientific ideas and engineering solutions with clarity through both the written and spoken word.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

21. When planning and teaching, educators teach students how mathematical concepts and skills apply to scientific investigation and engineering design.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

22. When planning and teaching, educators teach students content area reading skills that facilitate their ability to access, interpret, and evaluate scientific knowledge.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

Best Practices for Student Learning

23. When planning and teaching, educators have students ask questions about scientific phenomena that can drive exploration

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

24. When planning and teaching, educators focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

25. When planning and teaching, educators recognize that the development of student understandings of disciplinary core ideas is a progression that takes place over years.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

26. When planning and teaching, educators include core ideas that can be investigated at multiple grade levels with increasing sophistication.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

27. When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

28. When planning and teaching, educators use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

29. When planning and teaching, educators have students engage in sustained investigations accompanied by necessary teacher support.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

Challenging Crosscutting Concepts

30. When planning and teaching, educators have students develop an understanding that phenomena work differently at different scales.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

31. When planning and teaching, educators have students use systems thinking when investigating scientific phenomena.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

32. When planning and teaching, educators have students consider that since energy and matter are conserved, much can be determined by studying their flow into and out of systems.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

33. When planning and teaching, educators have students investigate phenomena in terms of structure and function as a means of sense making.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------

34. When planning and teaching, educators have students identify what aspects of a system remain stable over time and what aspects undergo patterns of change.

No Readiness	Slight Readiness	Fair Readiness	Solid Readiness	Strong Readiness	Advanced Readiness
--------------	------------------	----------------	-----------------	------------------	--------------------