

2009

Enhancing nature of science understanding, reflective judgment, and argumentation through socioscientific issues

Brendan E. Callahan
University of South Florida

Follow this and additional works at: <http://scholarcommons.usf.edu/etd>

 Part of the [American Studies Commons](#)

Scholar Commons Citation

Callahan, Brendan E., "Enhancing nature of science understanding, reflective judgment, and argumentation through socioscientific issues" (2009). *Graduate Theses and Dissertations*.
<http://scholarcommons.usf.edu/etd/1886>

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.

Enhancing Nature of Science Understanding, Reflective Judgment, and
Argumentation through Socioscientific Issues

by

Brendan E. Callahan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Secondary Education
College of Education
University of South Florida

Major Professor: Dana Zeidler, Ph.D.
Pat Daniel, Ph.D.
John Ferron, Ph.D.
Elaine Howes, Ph.D.

Date of Approval:
July 28, 2009

Keywords: discourse patterns, epistemology, moral and ethical issues, science education,
secondary education

© Copyright 2009, Brendan E. Callahan

DEDICATION

I dedicate my dissertation work to my family and friends at the University of South Florida. First, I would like to thank my mother, Jane Callahan for instilling a love of reading into me at an early age. I also dedicate this work to my loving and supporting wife, Tracy Callahan, and my two amazing children, Caitlin and John. Without your support and patience, this paper would have never been written. I would like to thank you Tracy, for not only helping me keep focus when the writing did not come as easily, but also for your technological assistance. Although I will never be able to express how large a role you played in the development of this dissertation, I plan to spend the rest of our lifetimes trying to express my gratitude.

I also thank the cohort of students who have travelled the road toward a doctoral degree with me. Your insights into science and education made each class enlightening, and working with all of you on various projects has provided me with the experience and knowledge needed to be successful in academia. To those of you who have completed, I look forward to the day we will reunite at conferences and gatherings. To those of you who continue to write I quote Louis Pasteur, “Let me tell you the secret that has led me to my goal: my strength lies solely in my tenacity.”

ACKNOWLEDGMENTS

I wish to thank my committee members who helped to generate this dissertation with their expertise and time. A special thanks to Sensei Dana “Surf Joe” Zeidler, my committee chair, for the number of hours he spent reflecting, refining, reading, and answering a multitude of questions from an unenlightened brute. The mentorship you have provided over the last nine years has given me a tremendous amount of insight into the academic life, and for this I will not be able to thank you enough. I would also like to thank Dr. Pat Daniel, Dr. John Ferron, and Dr. Elaine Howes for agreeing to serve on my committee. I appreciate all you have done for me in the development of this dissertation.

I would also like to thank my school district and high school for providing me the teaching experience I have developed over the last eight years as a public school teacher. The amount of knowledge I have gained from my time as a novice teacher has been immense, but still is never enough. The friends I have made throughout my time in the public school are dear to me and a constant source of inspiration. I would especially like to thank the science department at Dunedin High School. You are all amazing teachers, and although you do not hear it enough – you do make a difference!

Finally I would like to thank the administrators, teachers, and students that assisted me with this project. Your endless support and energy for this project made the research an enjoyable experience.

TABLE OF CONTENTS

LIST OF TABLES	v
ABSTRACT	vi
CHAPTER ONE: THE PROBLEM	1
Introduction	1
Background	2
Scientific Literacy	2
Socioscientific Issues and Impact on Learning	6
Argumentation	10
Research	11
Nature of Science	12
Reflective Judgment	12
Argumentation	13
Statement of Problem and Research Questions	15
Statement of Problem	15
Question 1 Rationale	16
Research Question 2	17
Question 2 Rationale	17
Research Question 3	18
Question 3 Rationale	18
Importance of the Study	18
Summary	19
CHAPTER TWO: LITERATURE REVIEW	21
Introduction	21
The Incorporation of Socioscientific Issues in the Curriculum	21
The Nature of Science in Science Education	24
Assessing Nature of Science Understanding	35
Views on Science and Education	35
Views on Nature of Science	38
Using Science to Develop Reflective Judgment	42
Assessing the Reflective Judgment Model	48
Reasoning about Complex Issues Test	48

Prototypic Reflective Judgment Interview	49
Argumentation is Central to Science and Society	51
Describing Argumentation	51
Argumentation should be Explicitly Taught in Science	57
Summary	58
CHAPTER THREE: METHODS	59
Research Design	59
Purpose	60
Research Questions	60
Population and Sample	60
Variables	63
Nature of Science	63
Reflective Judgment	63
Socioscientific Issues Curriculum	63
Instruments/Measures	63
Nature of Science	64
Views on Science and Education	64
Tentativeness of Scientific Knowledge	65
Nature of Scientific Knowledge	66
Nature and Comparison of Theories and Laws	67
Use of Imagination	68
Views on Nature of Science	69
Tentativeness of Scientific Knowledge	69
Nature of Scientific Knowledge	70
Nature and Comparison of Theories and Laws	70
Use of Imagination	71
Reflective Judgment	71
Argumentation	71
Written Argumentation	72
Oral Argumentation	73
Data Collection	74
Teacher Selection and Training	74
Graduate Student Training	78
Curriculum Development	79
Biology I Curriculum	84
Unit 1: Tokaimura Accident	84
Unit 2: Re-Enactment of the Kyoto Protocol	84
Unit 3: Stem Cells: Saving Superman	85
Unit 4: Transgenic Crops: Do You Really Know	85

What You're Eating?	
Unit 5: Pesticides: Can We Do Without Them?	86
Student Selection	86
Quantitative Procedures	89
Qualitative Procedures	91
Time Frame for Data Collection	93
Procedures for Maintaining Confidentiality	94
Data Analysis	95
Nature of Science	99
Reflective Judgment	100
Argumentation	100
Summary	101
CHAPTER FOUR: RESULTS	102
Introduction	102
Research Questions	102
Research Question 1	102
Tentativeness of Scientific Knowledge	105
Nature of Scientific Observations	109
Nature and Comparison of Theories and Laws	110
Use of Imagination in Science	113
Research Question 2	114
Reasoning about Complex Issues Test	115
Prototypic Reflective Judgment Interview	117
Research Question 3	119
Written Argumentation	120
Oral Argumentation	122
Summary of Results	124
CHAPTER FIVE: DISCUSSION	128
Introduction	128
Discussion of the Findings	128
Implications for Practice	136
Limitations of the Study	139
Recommendations for Further Research	141
Conclusions	141
REFERENCES	145
APPENDICES	155
Appendix A: Views on Science and Education survey	156

Appendix B: VNOS-B Questionnaire	160
Appendix C: RCI Protocol	161
Appendix D: Prototypic Reflective Judgment Interview protocol	165
Appendix E: Description of SSI Units Used in Study	167
Appendix F: Persuasive Essay Assignment	173
Appendix G: Persuasive Essay Assignment (posttest 2)	175
Appendix H: Rubric for analysis of oral argumentation	177
Appendix I: Argumentation Interview	178

ABOUT THE AUTHOR

End Page

LIST OF TABLES

Table 1	Consensus regarding the nature of science	26
Table 2	Congruence between VOSE and VNOS constructs	41
Table 3	Summary of reflective judgment stages	45
Table 4	Instruments used during the study	64
Table 5	Teacher schedules for first semester biology	75
Table 6	Correlation of SSI and Sunshine State Standards for Biology	77
Table 7	Description of student population: gender	87
Table 8	Description of student population: race and ethnicity	88
Table 9	Description of student population: age	89
Table 10	Data collection during the study	92
Table 11	Timeline for conducting study	94
Table 12	Pre and Post-test mean VOSE scores for nature of science constructs	103
Table 13	Pre and Post-test mean scores for the RCI	115
Table 14	Pre and Post-test mean argumentation essay scores	121

ENHANCING NATURE OF SCIENCE UNDERSTANDING, REFLECTIVE JUDGMENT, AND ARGUMENTATION THROUGH SOCIOSCIENTIFIC ISSUES

Brendan E. Callahan

ABSTRACT

There is a distinct divide between theory and practice in American science education. Research indicates that a constructivist philosophy, in which students construct their own knowledge, is conducive to learning, while in many cases teachers continue to present science in a more traditional manner. This study sought to explore possible relationships between a socioscientific issues based curriculum and three outcome variables: nature of science understanding, reflective judgment, and argumentation skill. Both quantitative and qualitative methods were used to examine both whole class differences as well as individual differences between the beginning and end of a semester of high school Biology I. Results indicated that the socioscientific issues based curriculum did not produce statistically significant changes over the course of one semester. However, the treatment group scored better on all three instruments than the comparison group. The small sample size may have contributed to the inability to find statistical significance in this study. The qualitative interviews did indicate that some students provided more sophisticated views on nature of science and reflective judgment, and were able to provide slightly more complex argumentation structures. Theoretical implications regarding the use of explicit use of socioscientific issues in the classroom are presented.

CHAPTER ONE: THE PROBLEM

Introduction

A divide exists between theory and practice in American education. While education researchers have presented arguments that science should involve more inquiry and analysis leading to conceptual change, many teachers continue to present science in a more traditional way. Consequently, what students learn in the classroom is disconnected from their daily lives (Duit & Treagust, 1998, National Research Council, 1996, 2000). The shortcomings of science education become apparent when students answer conceptual questions, or attempt to link science content to the real world. There are many aspects to scientific literacy, however, widespread consensus exists that the ability to use scientific concepts to solve new problems should be included in any definition of scientific literacy (American Academy for the Advancement of Science, 1989; Organisation for Economic Co-operation and Development, 1998).

Scientific literacy has not only been the focus of science education research, but also international assessments, such as the Programme for International Student Assessment (PISA), and Trends in International Mathematics and Science Study (TIMSS) (National Center for Education Statistics, 2004; Organisation for Economic Co-operation and Development, 2006). These standardized tests show students from the United States are falling behind other students from other countries in terms of scientific literacy. Students from the United States scored below the OECD average on the PISA 2006, and lower than would be predicted based on the gross domestic product per capita

for the country. Additionally, data from the 2003 TIMSS assessment revealed eighth grade students from the United States ranked ninth of forty-five countries on the science portion of the assessment. As these tests were aligned to more contemporary views of scientific literacy, evidence suggests that the science curriculum of the United States is not keeping pace with science education efforts around the world.

Within this framework, the overall goal of the study was to design, implement, and evaluate a semester-long high school biology curriculum aimed at enhancing students' understanding of three aspects related to scientific literacy: nature of science, reflective judgment, and argumentation skills. The curricular content was taught using a series of socioscientific issues to place biological content within real world applications.

Background

Scientific Literacy

The term “scientific literacy” has become unmanageable and difficult to succinctly define. A number of researchers and organizations have attempted to provide a holistic view of scientific literacy (AAAS, 1989; Hodson, 2003; OECD, 1998; Pella, O’Hearn & Gale, 1966). Norris and Phillips (2003) provided a comprehensive review of scientific literacy, as given from seventeen different research groups and major organizations. They found that the term scientific literacy has been used as:

- knowledge of substantive content of science and ability to distinguish science from non-science
- understanding science and its applications
- knowledge of what counts as science
- independence in learning science

- ability to think scientifically
- ability to use scientific knowledge in problem solving
- knowledge needed for intelligent participation in science-based social issues
- understanding the nature of science, including relationships with culture
- appreciation and comfort with science, including its wonder and curiosity
- knowledge of the risks and benefits of science
- ability to think critically about science and deal with the scientific enterprise (Norris & Phillips, 2003, p. 225)

Additionally, Norris and Phillips argued that the fundamental sense of the term “literacy,” the reading and writing of science, should be included as a core theme of scientific literacy. Zeidler (2007) argued that “any conception of what it means to be scientifically literate falls short of the mark if moral reasoning, ethical considerations, and an eye toward character are not part of our understanding of [scientific literacy](p. 1).”

Roberts (2007) outlined two visions of scientific literacy. Vision I related to science itself, particularly the “products and processes” (p. 730) of the scientific enterprise. Vision I included three concepts: basic scientific concepts, nature of science, and scientific ethics. Vision II related to the types of science students may encounter in the future. These situations will likely involve interactions between science and society. Three concepts related to Vision II included: “interrelationships of science and society, interrelationships of science and the humanities, and differences between science and technology.” (p. 739) The use of socioscientific issues directly relates to the Vision II concept of scientific literacy.

Researchers have used many formats to assess scientific literacy. National and international standardized tests such as PISA and TIMSS utilized the idea that scientific literacy involves the use of “scientific knowledge to identify questions and draw evidence-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity.” (OECD, 2003, p. 133) The OECD examined Bybee’s (1997) work in developing the goal of their assessment as “conceptual and procedural scientific literacy.” (p. 133) The PISA exam tests students’ knowledge of scientific concepts and processes within the framework of situations and contexts of science based issues. The TIMSS exam focuses on the scientific concepts needed by students, including the use of the scientific method to solve scientific questions; there is less emphasis on placing the questions within an issue-based context.

Additionally, science education researchers have investigated various aspects of scientific literacy. Iding and Klemm (2005) examined pre-service teachers’ ability to evaluate information from the World Wide Web. They found students used a variety of influences to determine the credibility of information found on websites. The authors promoted three recommendations for curriculum utilizing the Internet: that website evaluation should be explicitly taught throughout the curriculum in K-12 schools utilizing a variety of contexts, teachers should determine the cognitive load of their students, and researchers should develop a consensus for the criteria of website evaluation.

Kolstø, Bungum, Arnesen, Isnes, Kristensen, et al. (2006) examined students’ use of scientific information related to a SSI. The researchers found that students use three major criteria in determining the quality of argumentation: empirical and theoretical adequacy, completeness of information, and the social aspects of sources of information.

They found that incorporating science content knowledge was not sufficient for increased argumentation, which echoes Zohar and Nemet's (2002) research. Instead, science curricula should include the methodological norms, social processes, and institutional aspects of science in order to develop students capable of participating in SSI debates.

Sadler, Chambers and Zeidler (2004) examined high school students' use of nature of science concepts when dealing with a SSI, global warming. The researchers examined the aspects of empiricism, tentativeness, and social embeddedness of science during student interviews regarding the issue. One major result from the study was that forty percent of students were not able to identify and describe data. This finding has tremendous impact for both the use of a SSI based curriculum, which is heavily based on evidence and argumentation, as well as scientific literacy in general. The researchers also found that in-depth investigations were needed to confront students' core beliefs, that student relevance to issues and personal consequences often formed the basis of students' stances on an issue, and that forty percent of the students reported that the article that was more "scientifically meritorious" was less convincing, which showed that students utilize more than science content when examining an issue, which Kolstø (2001) termed content-transcending knowledge. He outlined a general framework of eight "content transcending" themes for examining the science dimension of SSI in science education that include: 1) Science-in-the-making and the role of consensus in science; 2) Science as one of several social domains; 3) Descriptive and normative statements; 4) Demands for underpinning evidence; 5) Scientific models as context-bound; 6) Scientific evidence; 7) Suspension of belief; and 8) Scrutinizing science-related knowledge claims.

Sadler and Donnelly (2006) noted that content knowledge was a necessary but not sufficient condition for learning argumentation in science. They argued that content knowledge was needed to engage in SSI, but that the content knowledge itself was not sufficient for argumentation regarding SSI. Students, on the other hand, also needed “context” knowledge, or in-depth study of the issues surrounding the content. A science course that blends the content and context aspects of science education has the potential to optimize the conditions for student learning.

Socioscientific Issues and Impact on Learning

The Socioscientific Issues (SSI) movement focuses on the incorporation of social issues involving a moral or ethical component with scientific relevance. Three main characteristics of the SSI movement are their open-endedness, their controversial nature, and the inclusion of moral or ethical reasoning (Zeidler & Sadler, 2008a; Zeidler, Sadler, Simmons & Howes, 2005). These components allow students to think critically on assigned issues, and discuss the topics with others who believe differently.

There are many influences on students’ thinking regarding socioscientific issues, including issues of personal experiences and perceived information quality (Sadler, Chambers, & Zeidler, 2004) and also by affective considerations (Sadler, 2002). With the SSI model students are forced to critically evaluate their own beliefs through social discourse and argumentation, resulting in the formation of personally relevant scientific knowledge.

There has been research to investigate student involvement with SSI over an extended period of time. These studies found that SSI were useful in confronting students’ core beliefs, connecting science to the real world, improving students’

understanding of scientific concepts, moral sensitivity, and improving students' reflective judgment (Fowler, Zeidler, & Sadler, 2009; Zeidler, Applebaum, & Sadler, 2006; Zeidler, Sadler, Applebaum & Callahan, 2009). It is important to note, however, that these studies were exploratory in nature and came from a single year-long treatment using multiple classes with one teacher. Confirming evidence with multiple teachers in different settings is needed to reinforce the arguments that SSI are useful in developing nature of science understanding and reflective judgment.

Nature of Science

Becoming scientifically literate involves having a contemporary view of the nature of science (NOS), which deviates from the belief that science is completely unattached and objective, to placing science within societal contexts. The American Academy for the Advancement of Science (1989) provided some of the first benchmarks regarding NOS teaching and learning, and the NRC (1996) expanded the discussion of NOS with a set of learning objectives at each level of K-12 education. McComas, Clough, and Almazroa (1998) examined the science standards for eight international organizations and compiled a comprehensive list of statements that summarized the nature of science. Their statements are organized here into the categories of scientific epistemology, scientific process, and history and sociology of science.

Scientific epistemology

- Science is an attempt to explain natural phenomena
- Scientific knowledge relies heavily, but not entirely, on observation, experimental evidence, rational evidence, and skepticism
- Science knowledge, while durable, has a tentative character

- Observations are theory-laden
- Laws and theories serve different roles in science, therefore students should note that theories do not become laws with additional evidence
- Science and technology impact each other

Scientific process

- There is no one way to do science (therefore, there is no universal step-by-step scientific method)
- Scientists are creative
- Scientists require accurate record keeping, peer review, and replicability
- New knowledge must be reported clearly and openly

History and sociology of science

- People from all cultures contribute to science
- Science is a part of social and cultural traditions
- History of science reveals evolutionary and revolutionary character
(McComas, Clough, & Almazroa, 1998, p. 6-7)

If scientific education is to be successful within this contemporary view of scientific literacy, all students must learn to think scientifically and understand the importance of science in their everyday lives (Kuhn, 1993). There is an understanding that we must educate the next generation of scientists and engineers, however, not every student is going to embark on a scientific career. All students will be involved in making scientific decisions in the future for their own families, and must have some

understanding of the science that affects their lives on a daily basis (Davies, 2004; Hodson, 2003; Symington & Tytler, 2004).

More recently, Lederman (2007) reviewed NOS research from the past fifty years. Although the researchers and instruments have changed over time, the theme that both students' and teachers' views of NOS are inadequate have remained constant. Some pedagogical considerations involve the use of "explicit, reflective instruction" (p. 869) rather than expecting NOS views to change through implicit instruction. Also, Lederman noted that science teachers did not regard NOS concepts as highly as traditional subject matter concepts. This pervasive view, as a result, perpetuated the belief by students that science does not relate to their everyday lives.

Reflective Judgment

Scientific literacy involves the ability to think critically regarding scientific issues. One epistemological model, the Reflective Judgment Model (RJM), examines people's views of knowledge over time. The RJM was developed by King and Kitchener over twenty years ago and has been supported by two decades of further research (King & Kitchener, 1994, 2002; Kitchener, 1983; Kitchener, King, Wood & Davison, 1989). They suggested that peoples' view of knowledge change in developmental stages over time starting in young adolescence and continuing through adulthood. The RJM is similar to other developmental models (Broughton, 1978; Fischer, 1980; Perry, 1970; Piaget & Inhelder, 1969) in the organization and order of stages, the consistency of responses across subjects at a given level, the discrete thought processes across stages, and hierarchical integration among stages. Reflective judgment involves the reasoning

patterns individuals use to support their position to ill-structured problems (King & Kitchener, 1994).

The reflective judgment model and socioscientific issues movement are in many ways analogous as both utilize ill-structured problems, focus on evidence and the analysis of positions, and examine problems with moral or ethical components. Reflective thinkers are able to synthesize multiple lines of evidence to arrive at a “best” resolution, and SSI scenarios often involve reasoning based on economic, political, moral, and cultural concerns, in addition to scientific reasoning. As this is the type of reasoning that occurs in the “real world of dirty sinks, and messy reasoning” (Zeidler & Keefer, 2003, p. 8), it is useful to involve students in the type of problem solving they will be exposed to as voting citizens, as politicians, or as scientists. Instruction, therefore, should involve the goal of allowing the student to become more sophisticated in reasoning, judgment, and debate through the use of increasingly complex sociomoral issues (Berkowitz, Oser & Althof, 1987; King & Kitchener, 1994).

Argumentation

The production of students capable of utilizing argument regarding scientific issues is one of the purposes of education (Driver, Newton, & Osborne, 2000; Ferretti, Andrews-Weckerly & Lewis, 2007; Sadler, 2006). The argumentative process is utilized in the scientific domain regarding the analysis of data, as well as in society regarding the outcome of science-based issues. The development of critical thinking skills depends upon the use of argumentation within the science curriculum (Shakirova, 2007; Waghid, 2005). Despite the importance of argumentation, studies have shown argumentation is not emphasized in classrooms (Newton, Driver, & Osborne, 1999) or well developed in

students (Sandoval & Millwood, 2005). Adolescents can develop more advanced argumentation patterns through practice and knowledge of argumentation norms (Felton, 2004; Walker & Zeidler, 2007; Weinstock, Neuman, & Tabak, 2004).

A SSI curriculum develops argumentation skills through the use of classroom debates and social negotiation of issues. As science is interwoven with such fields as medicine, law, economics, ethics, and politics, each of these areas influence student thinking regarding science, and should be examined within the context of scientific dilemmas. A SSI curriculum enables science content area teachers to offer a multidisciplinary approach to learning science, which provides a realistic context for science as well as a method for developing conceptual awareness in the science classroom (Sadler, 2006). As science is an integral part of society, it is important for students to examine the scientific enterprise through the multiple lenses characteristic of contemporary society. Additionally, a SSI curriculum provides the multi-faceted approach to science necessary for students to develop the critical thinking and literacy skills needed to succeed in an increasingly complex society.

Research

The goal of this study was to examine the potential benefits of utilizing SSI in the science curriculum by analyzing the effects of an SSI based curriculum in high school biology classes on three desired outcomes of science education: nature of science understanding, development of reflective judgment, and written argumentation. These areas have importance for science education due to their impact on future society.

Nature of Science

Nature of science understanding provides the background needed for future citizens to provide input regarding scientific decisions, as scientists, citizens, or politicians. The questions on the PISA exam are contextual-based, including questions based on issues that adults and policy makers encounter. The writers of the test require students to identify scientific issues, explain scientific phenomena, and use scientific evidence. Knowledge about the natural world (scientific themes) and how science works (nature of science) are needed as well in order to be successful on the test.

The reciprocal relationship between nature of science understanding and resolution of socioscientific issues has been discussed (Abd-El-Khalick, 2003; Dotger & Jones, 2007; Zeidler et al., 2009). Abd-El-Khalick found that college students had difficulty in transferring NOS understanding to the utilization of scientific knowledge of a socioscientific scenario, while both Abd-El-Khalick and Zeidler et al. have discussed that successful discussion of an SSI involves utilizing an informed view of the nature of science. To date, there have not been any long-term studies to empirically gauge the relationship between an SSI based curriculum and changes in nature of science understanding.

Reflective Judgment

Reflective judgment, which has parallels to higher order thinking, is an epistemological construct designed to examine how people conceptualize and acquire knowledge. Reflective thinkers possess the ability to analyze multiple lines of evidence and combine them to determine an outcome. In a SSI curriculum, these outcomes are socially negotiated in the classroom through argumentation and consensus-building.

Scientific literacy has been defined, in part, as "... the capacity to use scientific knowledge, to identify questions, and to draw evidence-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity (OECD, 2006, p. 133)." This conceptualization of scientific literacy is reflected in the PISA assessment by the use of questions framed within contemporary issues to assess student understanding of science. These questions require higher levels of reflective judgment, as they involve multiple lines of evidence, rather than the typical objective, fact-based science questions often asked in American classrooms.

One year long study investigated the link between reflective judgment and SSI in high school students, involving anatomy and physiology students (Zeidler et al., 2009). This study expands the literature base in this area by providing a slightly younger (Biology I) students as well as an analysis of the Reasoning about Current Issues (RCI) Test, a computer based test which was used in concert with the semi-structured interview process of the Prototypic Reflective Judgment Interview (PRJI). The previous study involved four classes taught by the same instructor, this study will use multiple instructors and classes.

Argumentation

Standardized exams frequently document students' lack of writing ability. The most recent NAEP results (2002) show that 26% of twelfth grade students wrote below "basic" achievement levels, while another 50% possessed "basic" level writing ability. Only 24% of students wrote at "proficient" or "advanced" levels (Education USA, 2003). As a society dependent upon the transmission of ideas, the lack of writing ability is

disturbing. As students progress into secondary and post-secondary education, there is less emphasis on narrative (story telling) and more emphasis on expository (descriptive) and persuasive writing. Further, as science advances through intellectual discourse, the inclusion of argument should be a part of the curriculum in each science class.

Theoretical evidence regarding the possibility of SSI contributing to improvement in argumentation can be inferred from previous work examining the necessary relationship between fundamental literacy and scientific literacy (Norris & Phillips, 2003). They argue since western science depends upon text as a medium, fundamental literacy is necessary to learn science. Studies (Sandoval & Millwood, 2005; Sippess, 2004) have found that students often have difficulty in relating evidence to claims across content areas, and students tend to exclude nature of science understanding (Walker & Zeidler, 2007) and content knowledge (Sadler & Donnelly, 2006; Sadler & Fowler, 2006) during debate. Long term treatments involving many opportunities to utilize argumentation skills may provide students the experiences needed to fully engage in argumentation. These long term treatments have not been studied in detail.

The use of a SSI based curriculum as the context for science curriculum may provide benefits not typically seen in short term treatments. Much of the literature base examines outcomes after a single SSI unit, however, given the complexities of nature of science understanding, the development of reflective judgment, and developing complex argumentation patterns, a longer treatment period may be more effective in revealing potential benefits to a SSI based curriculum. Zeidler and Sadler (2008b) argued students needed to develop a personal and relevant relationship with scientific issues through active participation in developing argumentation skills, the ability to distinguish between

science and non-science, and the evaluation of evidence and data. These are skills that develop over time, therefore, a longer treatment may be more useful in understanding the changes that occur in these areas over time.

Statement of Problem and Research Questions

Statement of Problem

Although the SSI field is rich with theoretical literature (Sadler, 2004; Zeidler et al., 2005; Zeidler, Walker, Ackett, & Simmons, 2002), there has been relatively little empirical research to determine to what extent the inclusion of SSI develops nature of science understanding (Zeidler et al., 2002), promotes reflective judgment (Callahan, Zeidler, Cone & Burek, 2005; Zeidler et al., 2009), and increases argumentation (Zohar & Nemet, 2002), particularly over an extended period of time. The current study closed some of the gaps in the empirical literature by investigating three main areas: evaluate to what extent the inclusion of SSI into a semester-long curriculum provides more sophisticated views in students' nature of science understanding (Bell & Lederman, 2003), determine the extent to which SSI can be used to develop reflective judgment in high school students (Callahan et al.; Zeidler et al.), and evaluate to what extent the inclusion of SSI provides changes in students' argumentation skills. As these outcomes of science education are distinct, the following research questions were asked:

Research Question 1

To what extent do students enrolled in high school science classes utilizing socioscientific issues show greater development in nature of science understanding compared to students in traditional high school science classes?

Question 1 Rationale

A gap exists between science education theory and educational practice in the high school classrooms. Researchers have provided a thorough examination of the contemporary nature of science (Abd-El-Khalick, Bell & Lederman, 1998; Lederman, Abd-El-Khalick, Bell & Schwartz, 2002; McComas, Clough & Almazroa, 1998) in academic journals, however, many high school textbooks discuss scientific method and nature of science in chapter one, and fail to continue the themes throughout the course.

Alternatively, this course with multiple socioscientific issues involved the re-examination of the nature of science as it related to contemporary society throughout the course. Quantitative data was gathered by using the Views on Science and Education (VOSE, Chen, 2006a). This survey has been given to Taiwanese pre-service teachers in order to elicit their views on the nature of science. The concepts tested on the VOSE correlated well with previous NOS instruments, including the Views on Nature of Science (VNOS) surveys. Pretest and posttest administration of the VOSE provided the empirical data needed to gauge the treatment's effectiveness. Additional data was gathered by interviewing a subsample of the students following the posttest administration of the VOSE. These interviews focused on student changes in NOS over the course of the semester.

Research Question 2

To what extent do students enrolled in high school science classes utilizing socioscientific issues show greater development in reflective judgment compared to students in traditional high school science classes?

Question 2 Rationale

Science is often simply defined as the study of nature. However, science entails a characteristic mindset, with an emphasis on asking questions, collecting and analyzing data, and combining multiple perspectives into a cohesive whole that is used in many areas, including business, law, and the social sciences. The understanding and development of students' epistemological growth is vitally important for the continuation of a democratic society (Davies, 2004; Hodson, 2003; Symington & Tytler, 2004). Past work (Callahan et al., 2005; Zeidler et al., 2009), suggested that a socioscientific issues based curriculum can improve reflective judgment over time. However, these conclusions should be considered preliminary due to the limited sample size necessitated by the use of an interview process. There was a possibility that greater significance or more generalizable results would be found by the use of a computer-based survey, which was administered to a large group of students across multiple classrooms. Previous research utilized the Prototypic Reflective Judgment Interview protocol, which, due to the time consuming nature of interview protocols, necessitated the use of a smaller sample size ($n = 40$) distributed across four classes than could be measured utilizing a computer-generated survey. The use of multiple classrooms and a larger number of students provided additional evidence regarding the efficacy of a SSI based curriculum on reflective judgment.

Research Question 3

To what extent do students who are enrolled in high school science classes utilizing socioscientific issues show greater argumentation skills compared to students in traditional high school science classes?

Question 3 Rationale

Many science education researchers and organizations have identified argumentation as one of the core processes all students should develop during their K-12 years (AAAS, 1993; Driver, Newton & Osborne, 2000; NRC, 1996; Sandoval & Millwood, 2005) due to the central role of argumentation in producing scientific knowledge. Additionally, researchers (Ferretti et al., 2007; Sadler, 2004; Sadler & Donnelly, 2006; Shakirova, 2007; Walker & Zeidler, 2007) have indicated that the development of argumentation must involve socioscientific issues to promote moral and ethical reasoning along with citizenship education. Research suggested that secondary students have not developed mature argumentative reasoning (Felton, 2004; Marttunen, Laurinen, Litosselitti, & Lund, 2005; Watson, Swain, & McRobbie, 2004), and that many teachers do not incorporate argumentation in the science curriculum (Newton, Driver & Osborne, 1999). The development of teachers' knowledge and ability to teach argumentation must be a consideration when developing a curriculum based on socioscientific issues.

Importance of the Study

This study had the potential for both practical and theoretical significance. The main practical outcome was the development of a socioscientific issues based curriculum which could be used by teachers in the high school classroom setting. The process of training teachers to implement a socioscientific issues based curriculum, while not directly part of the study, also provided information regarding the issue teacher expertise and experience when dealing with the incorporation of a SSI curriculum. This

information is valuable to progress SSI from the domain of science education researchers to practicing teachers.

The theoretical importance related to providing empirical evidence regarding the utility of socioscientific issues in the science curriculum. A socioscientific issues based curriculum has the potential to have implications for such areas as nature of science, epistemology, character education, and literacy, although the long-term effects of SSI treatments are not well documented. Additionally, data collected through the use of the RCI tests provided significant information to the field of science education through the testing of high school students, as the RCI has not been widely used with high school students (Summers-Thompson, personal communication, March 2008). The use of a large heterogeneous sample provided evidence regarding the outcome variables directly measured in this study, as well as two assessments commonly used in science education.

Summary

Scientific literacy has been a major focus of science education for the past twenty years with the publication of such documents as *Project 2061: Science for All Americans* (AAAS, 1989) and the *National Science Education Standards* (NRC, 1996) and international tests such as the Trends in International Mathematics and Science Study and Programme for International Student Assessment (National Center for Education Statistics, 2004; Organisation for Economic Co-operation and Development, 2006). Although there is much debate regarding the definition of scientific literacy, many agree that an understanding of scientific processes and knowledge as well as the relationship between science and society are key elements of scientific literacy. Within this framework, a socioscientific issues curriculum was developed for high school Biology I

students in order to assess the development of three aspects related to scientific literacy: nature of science, reflective judgment, and argumentation. As the SSI movement seeks to investigate scientific issues with a moral or ethical component, it was reasoned that the discussion and debate regarding contemporary issues would facilitate students' development of scientific literacy.

CHAPTER TWO: LITERATURE REVIEW

Introduction

The central argument underlying the theoretical framework is that socioscientific issues have positive effects on students' scientific literacy in the classroom, and these effects need to be explored in more detail. To this end, a brief introduction to the framework guiding SSI instruction will precede arguments providing evidence that the use of a SSI-based curriculum would be beneficial to understanding the outcome variables under consideration: nature of science, reflective judgment, and argumentation. Links to both scientific literacy and SSI will be included in the discussion of the outcome variables to show the connection between scientific literacy, socioscientific issues, and the outcome variables. Background research will be presented to describe each of the outcome variables, as well as research involving prior methods of assessment and factors influencing the development of these traits and skills. Connections between nature of science, reflective judgment, and argumentation with SSI will justify the incorporation of SSI as a means for examining these outcomes.

The Incorporation of Socioscientific Issues in the Curriculum

The Socioscientific Issues (SSI) movement on the surface focuses on the incorporation of science issues with social relevance. Current issues include: genetic engineering, cloning, stem cell research, and alternative fuel sources. Some characteristics of socioscientific issues include the belief that there are no easy or correct answers for the issues involved, these issues need to be personally relevant to students,

and there must be a moral or ethical component to them (Zeidler et al., 2005). These characteristics allow students to engage in these issues, and then discuss their viewpoints with students who believe differently. Driver, Newton, & Osborne (2000) stated that the purpose of science education should be to analyze arguments relating to the social application of scientific issues.

The Socioscientific Issues (SSI) movement evolved from the belief that the previous STS (science-technology-society) model, while connecting science to real life, did not adequately address the moral development and character education components to critically evaluate scientific dilemmas (Zeidler et al., 2005). Some of the problems with the STS model include: the marginalization of STS education, the lack of personal relevance, the lack of an ethical component as a foundation, the possibility of teaching about issues, rather than students engaging in issues, and the lack of a unifying theoretical framework. The SSI model utilizes argumentation and debate to challenge students' personal beliefs, resulting in socially negotiated scientific knowledge. As the development of scientific knowledge is attained through discourse and analysis, the utilization of an SSI curriculum provides students with the thinking skills needed to develop scientific literacy, and many educators have argued for the inclusion of SSI in the science curriculum (Driver, Newton & Osborne, 2000; Kolstø, 2001; Sadler, 2004; Zeidler & Keefer, 2003; Zeidler et al.; Zeidler & Sadler, 2008a).

A number of papers have focused on some of the main attributes of socioscientific issues, including the role of affect in moral matters (Sadler, 2002), the effect of critical thinking skills on nature of science understanding regarding SSI (Sadler, Chambers, & Zeidler, 2004), and reflective judgment (Callahan et al., 2005; Zeidler et al., 2009). SSI

have also been shown to increase content knowledge due to the personal relevance in students (Zeidler, Applebaum, & Sadler, 2006). Students who engage in transactive peer discussions, discussions in which an individual “transforms” the reasoning of another, by elaboration, critique, extension, or integration with one’s own reasoning, solve both scientific and mathematical problems more effectively than students who had fewer transactive peer interactions (Berkowitz & Simmons, 2003). This research has implications for utilizing a socioscientific issues-based curriculum, which by definition involves transactive discussions through the process of argumentation. To realize this goal, however, teachers must establish classroom environments designed to promote the safe expression and evaluation of ideas (Zeidler & Sadler, 2008b).

One study (Callahan, 2009) investigated teacher and student beliefs regarding the skills students would need to be successful in engaging in a socioscientific issues based curriculum. Interviews with two teachers and twenty four students involved in a semester long socioscientific issues based Biology curriculum at a single public high school in the Tampa Bay area provided themes regarding what students would need to engage in a socioscientific issues based curriculum. Teachers and students responded that literacy skills, and research/information skills were sometimes lacking, and students also responded that classroom dynamics also plays a role in learning science through a socioscientific issues based curriculum. The study found that explicit attention to the skills learned during the research and debate processes within the framework of a socioscientific issues based curriculum has the potential to provide experiences designed to increase students’ fundamental and informational literacy.

The Nature of Science in Science Education

Contemporary science educators and organizations (AAAS, 1989; Abd-El-Khalick, 2005; Abd-El-Khalick & Lederman, 2000; Holbrook & Rannikmae, 2007; NRC, 1996; Schwartz, Lederman, & Crawford, 2004, among others) have argued that school science should emphasize nature of science. They argue that study of the nature of science places science within societal context, which comprises a fundamental aspect of scientific literacy. There has been much debate regarding the teaching of the nature of science, and while there is discussion regarding the boundaries of the concept, there appears to be some agreement on the main concepts of the nature of science. Lederman et al. (2002) provided theoretical backing for the development of the Views on Nature of Science questionnaire by describing aspects of the nature of science most relevant to K-12 education: that scientific knowledge is tentative, empirical, theory-laden, and product of human inference, involves imagination and creativity, socially and culturally embedded, the distinction between observation and inference, lacks a universal scientific method, and the relationship between theories and laws. McComas, Clough, & Almazroa (1998) examined the nature of science explanations from eight international science education standards documents. They found a variety of concepts that signify nature of science understanding. Their statements can be divided into scientific processes, the history and sociology of science, and science epistemology. Scientific process can be thought of as how scientists conduct scientific research. In studying scientific process, the skills of questioning, developing novel procedures, gathering and analyzing data creatively, and reporting results are examined. The history and sociology of science involves an understanding that scientific knowledge takes place within a societal context,

and that people's beliefs and priorities often determine areas of scientific research. Science epistemology involves science as knowledge, or the definition of science. Osborne, et al. (2003) used a three part "Delphi study" to examine the level of agreement between a wide range of experts within the science community. The Delphi panel included five scientists, five historians, philosophers, and sociologists of science, five science educators, four science teachers, and four science communicators. The three part process culminated in nine themes that were determined by consensus (two-thirds or more rated the theme a four or higher on a five point Likert scale) and by stability (less than one-third of the members changed their ratings between rounds two and three). The nine themes map well onto previous work by Lederman et al. and McComas et al., and reflected in national science standards (AAAS, 1989). Table 1 shows this congruence of thought.

Table 1

Consensus Regarding the Nature of Science

Concept	Lederman et al, 2002	McComas, et al., 1998 (p. 6-7)	Osborne, et al., 2003	AAAS (1989)
Scientific epistemology	<p>Science is partially based on observations of the natural world. Students should be able to distinguish between observation and inference (p.500). Knowledge, although reliable and durable, is never absolute or certain. (p. 502)</p> <p>Knowledge is theory laden.</p> <p>There is a distinction between theory and law.</p>	<p>Science is an attempt to explain natural phenomena</p> <p>Scientific knowledge depends on observation, experimentation, evidence, and skepticism</p> <p>Knowledge is durable, yet tentative.</p> <p>Observations are theory-laden</p> <p>Laws and theories serve different roles.</p>	<p>Hypothesis and Prediction</p> <p>Analysis and Interpretation of Data</p> <p>Science and Certainty</p> <p>Science and Technology</p> <p>Creativity Science and Questioning</p> <p>Scientific Method and Critical Testing</p> <p>Cooperation and Collaboration</p> <p>Cooperation and Collaboration</p> <p>Historical Develop. of Sci. Know. Diversity of Sci. Thinking</p>	<p>Occasionally major shifts occur in the view of how the world works, usually knowledge occurs in small advances. The testing, revising, and occasional discarding of theories never ends.</p> <p>Basic beliefs about the value of evidence, logic, and good arguments.</p>
Scientific Process	<p>Myth of the scientific method. Generating scientific knowledge also involves imagination and creativity. (p. 500)</p>	<p>Science and technology impact each other</p> <p>There is no one way to do science</p> <p>Scientists are creative.</p> <p>Scientists require accurate record keeping, peer review, and replicability.</p>	<p>Science and Technology</p> <p>Creativity Science and Questioning</p> <p>Scientific Method and Critical Testing</p>	
History and Sociology of Science	<p>Social and cultural embeddedness of scientific knowledge</p>	<p>People from all cultures contribute to science.</p> <p>Science is a part of social and cultural traditions.</p> <p>History of science reveals evolutionary and revolutionary character.</p>	<p>Cooperation and Collaboration</p> <p>Cooperation and Collaboration</p> <p>Historical Develop. of Sci. Know. Diversity of Sci. Thinking</p>	

Incorporating the epistemology of science into the curriculum indicates a shift from an absolute view of science to science within a societal context, leading to a contemporary view of scientific literacy. In addition to the major concepts that form the core of scientific knowledge, students are expected to have an understanding of the processes scientists use to accumulate knowledge, as well as the nature of scientific knowledge (AAAS, 1989; NRC 1996). The transition from an empirically based nature of science to a societal based nature of science coincided with Kuhn's (1962) *Structure of Scientific Revolutions* (Giere, 1988, from Abd-El-Khalick & Lederman, 2000). A contemporary nature of science understanding also diverges from traditional science education in that whereas the traditional form of science education focused on the transmission of facts, contemporary nature of science deals more with the thinking processes of science. Rudolph (2003) utilized historical case studies of John Dewey and Joseph Schwab as support of the premise that political and societal factors inform science epistemology at any given point in time. The contemporary nature of science can be traced to Dewey, in that process is favored over fact. "Instruction in scientific thinking, not science per se should be the primary aim of the science teacher." (p. 69) He also mentions Dewey's affinity for science and the scientific process when he quotes, "The intellectual practices of science were to serve as the model for rational thought in all affairs regardless of their domain." (p. 69)

If one accepts the premise that scientific thought is beneficial in a variety of domains, it follows that all students must be trained in scientific processes, regardless of their future. Deanna Kuhn (1993) states this when she writes,

Scientific thinking tends to be compartmentalized, viewed as relevant and accessible only to the narrow segment of the population who pursue scientific careers. If scientific education is to be successful, it is essential to counter this view, establishing the place that scientific thinking has in the lives of all students. A typical approach to this objective has been to try to connect the *content* of science to phenomena familiar in students' everyday lives. An ultimately more powerful approach may be to connect the *process* of science to thinking processes that figure in ordinary people's lives. (p. 333)

The idea that science education prepares students for roles as both practicing scientists and democratic citizens capable of decision making is not new (AAAS 1989, 1993; Hodson, 2003; Kolstø, 2001; McComas & Olson, 1998; Symington & Tytler, 2004), however, the incorporation of discussion and argumentation has not been fully embraced by classroom teachers (Driver, et al., 2000). They found many students were taught from a positivist perspective, leading to the idea that scientific knowledge is factual and unchanging. Perhaps this perception stemmed from the mode of classroom delivery, namely teacher talk. The researchers performed an analysis of thirty four science lessons in the United Kingdom involving eleven to sixteen year old students. They found that in only two instances there were student discussions, and each of them lasted less than ten minutes. When the classroom culture is one of sitting and listening to the teacher, it follows that the students are going to view the teacher as the authority in the field, not to be questioned.

Nature of science understanding provides the background needed for future citizens to provide input regarding scientific decisions, as scientists, citizens, or politicians. The questions on the PISA exam are contextual-based, including based on issues that adults and policy makers encounter. The writers of the test require students to identify scientific issues, explain scientific phenomena, and use scientific evidence. Knowledge about the natural world (scientific themes) and how science works (nature of science) are needed as well by all students as scientific thinking forms the basis for many disciplines, as well as the similarities between scientific thinking and informal reasoning.

Once the case has been made that nature of science should be taught to all students, a logical step would be to investigate the question of NOS understanding in the classroom setting. Research (Abd-El-Khalick, 2005; Abd-El-Khalick & Lederman, 2000) shows that implicit instruction does not improve NOS understanding. Implicit instruction involves the nature of science being learned during the course of normal science coursework, some process instruction, and laboratory work.

An alternative to implicit NOS instruction is explicit instruction. Explicit instruction in nature of science involves direct instruction of NOS concepts, including the history and philosophy of science. The history and philosophy aspects of science allow students to understand the role of science within the social and political culture in which it operates. There has been some research to determine what approaches are beneficial to promoting NOS understanding. Abd-El-Khalick (2005) investigated the use of a philosophy of science course with pre-service secondary teachers. The philosophy of science course was the third in a sequence, which some students took along with Science Methods II. The first course in the sequence, Science Methods I, was designed to give

the students basic background regarding teaching, as well as explicit instruction in NOS. Science Methods II was a course designed to give the pre-service teachers a model for teaching in their own classrooms, with an emphasis on inquiry activities to provide experiences to be replicated with their own students. This course also emphasized the planning of science lessons, utilizing media and other resources, and applying various approaches to teaching science. The philosophy of science course focused on issues relevant to science education. The students wrote four extended reflection papers during the semester course. The central research questions were: did the teachers gain understanding of the target aspects of NOS, how did the teachers perceive teaching NOS in the future, and did they develop instructional planning related to NOS concepts. The researcher found that although both groups (methods classes only vs. methods and philosophy of science) did gain some knowledge regarding NOS concepts, many of the methods only group held naïve conceptions of NOS as tested by the VNOS Form C. On the other hand, all of the students in the philosophy of science course had developed informed views of NOS as tested on the VNOS-C. The researcher also found there were a large number (34%) of the methods students who were hesitant to present science as complicated and without uniform thought, and 25% of the students thought NOS should be taught, but that it was not possible to teach the subject to secondary students. The pattern was similar for the philosophy of science students for their first two reflection papers, after which the researcher noticed a shift that diverged from the methods students. The third and fourth reflection papers tended to show more of an interest in incorporating reflection, behaviors, and assignments indicative of a willingness to teach NOS.

There is also evidence to suggest that explicit instruction helps high school students gain a more sophisticated understanding of NOS. Khishfe & Lederman (2006) examined the use of integrated and non-integrated forms of explicit instruction in NOS with ninth grade environmental students. They utilized two intact classes in which all students completed a nature of science questionnaire that was developed by the researchers. Ten students, five from each class, were then selected for interviews. A six week treatment followed, with the integrated group using a unit on global warming, while the control group studied NOS concepts independent of science content. The researchers found that both groups improved NOS understanding at the end of the six weeks, with the non-integrated group performing slightly better at moving from naïve to transitional views, and the integrated group performing slightly better at moving from transitional to informed views of NOS.

Both of these studies make the case that the nature of science should be explicitly taught if we expect all our students to have an informed view of NOS. Based on the research to date, classroom teachers should incorporate more non-integrated instruction of NOS at the beginning of the year, and transition to integrated instruction as the course progresses. Additionally, the fields of history and philosophy of science have been shown to improve NOS understanding in young adults. The next step then, is to develop a curriculum that is able to accommodate these varying concepts of integrated explicit instruction while incorporating the history and philosophy of science in a way that is relevant to today's students. The answer lies in the use of socioscientific issues as the basis for science instruction.

The reciprocal relationship between nature of science understanding and resolution of socioscientific issues has been discussed by some researchers (Abd-El-Khalick, 2003; Dotger & Jones, 2007; Zeidler et al., 2009), and in a few instances been investigated in small units (Bell & Lederman, 2003; Sadler, Chambers & Zeidler, 2004; Walker & Zeidler, 2007). Abd-El-Khalick found that college students had difficulty in transferring NOS understanding to the utilization of scientific knowledge of a socioscientific scenario, while both Abd-El-Khalick and Zeidler et al. have discussed that successful discussion of an SSI involves utilizing an informed view of the nature of science.

Bell and Lederman (2003) sought to study the influence of NOS understanding on science and technology based decision making processes. They chose to study adults, primarily due to the inability to find secondary students with mature conceptions of NOS, as well as the probability that the adults were more likely to have made some decisions regarding science and technology issues. They gathered two groups, one considered NOS experts, and the other NOS novices. Both groups consisted of university professors, however, the experts were taken from the fields of science education, science philosophy, and research science. The novices were sampled from the humanities and business. Both groups were given the Decision Making Questionnaire (DMQ), which the authors produced in order to provide open ended answers to four science and technology based scenarios (fetal tissue usage, global warming, diet and exercise, and smoking/cancer link). The subjects were subsequently interviewed using the information from the DMQ to guide the interviews to construct profiles of each of the subjects. Each of the subjects also completed the VNOS Form B questionnaire. The researchers found little difference

in the use of NOS concepts when involved with decision making on scientific and technology issues. Rather, the groups tended to use a variety of influences (moral and ethical values, social, political, pragmatism) to reach decisions on the topics. The NOS expert group did tend to look at the evidence more critically and in combination with other reasoning patterns, while the NOS novice group tended to examine evidence as proof, and became skeptical of the scientific process when an absolute truth was not offered.

Sadler, Chambers, & Zeidler (2004) investigated NOS concepts in the context of a socioscientific issue, global warming. They investigated students' ability to interpret data, and looked at the NOS concepts of cultural embeddedness and tentativeness. The researchers also examined the students' ability to interpret and evaluate conflicting data. They gathered two intact classes of high school biology classes and gave each student a fictitious "science brief" on global warming. Although the brief was developed by the researchers, the brief was designed to present parallel data taken from contemporary arguments regarding global warming. Each student was then given an open-ended questionnaire to complete. From this group, a subsample of students was interviewed regarding their feelings. The selection was purposeful, as the authors sought to involve students with a range of critical thinking abilities for the interviews. The authors developed a hierarchy of data, ranging from "data confusion" or a misinterpretation of the data, to "data recognition, description, and explanation." The researchers found a variety of social influences guided students' thought, including economics, personal perspectives, societal causes, and societal effects. Some of the major findings include the alarming statistic that one half of the students were not able to identify and describe data

within the context of the issue. The authors also found that students did understand the use of societal factors in science as well as the tentative nature of scientific knowledge.

Walker & Zeidler (2007) investigated the use of scaffolded inquiry to assess argumentation and NOS conceptions on a SSI, genetically modified foods. They utilized two high school classrooms taught by the same teacher. One intact class was the control group, which was based on traditional “textbook” learning, while the treatment group studied genetically modified foods for a total of seven blocks of one and a half hours each. The treatment group utilized an online curricular unit to scaffold information. The researchers found that although the students were able to conceptualize NOS concepts, they did not utilize them effectively during the classroom debates, rather the students tended to use affective, emotional, and personal knowledge to support their positions. The researchers argued that future research should involve the incorporation of SSI into content units, as well as the practicality and effectiveness of a course utilizing the combined SSI-content curriculum over a year. The present study attempted to address those concerns.

Each of these studies investigated the use of NOS conceptions when dealing with SSI or decision making tasks. Although two studies (Sadler, Chambers, & Zeidler, 2004; Walker & Zeidler, 2007) both used biology based classes of high school students, neither one utilized multiple classrooms across multiple teachers, and neither study investigated the use of SSI over a long-term treatment. Perhaps the immersion in scientific thinking through the consistent use of SSI will train students to incorporate more scientific thinking in decision making scenarios.

Assessing Nature of Science Understanding

Throughout the last century, the development of an informed view of the nature of science has been a major goal of science education (Lederman, et al., 2002), and there have been many assessments designed to measure students' views of the nature of science as described by Lederman, Wade, & Bell (1998). However, a major argument against the use of these instruments was that the researchers and respondents viewed the questions differently. Aikenhead, Ryan, & Fleming (1989) developed the Views on Science-Technology-Society (VOSTS) utilizing a unique process. They initially used a free response format in order to gather data from the respondents. The subsequent multiple choice survey contained a range of responses generated from the open ended version of the survey. Both of the surveys used in this study, the VNOS and the VOSE, were developed using a similar process of a free response survey followed by interviews with the respondents to clarify ambiguous responses on the survey.

Views on Science and Education

The VOSE was written in response to criticisms with both forced-choice questionnaires such as the VOSTS and open ended surveys such as VNOS. Chen (2006a) found that the oversimplification of the questions in the VOSTS created ambiguity between the respondents and the researchers. Similarly, Chen argued that some concepts of NOS, such as scientific method, were not asked specifically, rather they were embedded in other questions. The lack of explicit focus could have potentially created ambiguity between respondents and researchers. She also noted that the VNOS was a challenging test for respondents to complete in 40-60 minutes.

The development of the VOSE was conducted in three phases. The first phase involved a review of the NOS literature and a pilot study with college students regarding NOS and teaching attitudes. Seven concepts of NOS were examined:

1. Tentativeness of scientific knowledge
2. Nature of observation
3. Scientific methods
4. Hypotheses, laws, and theories
5. Imagination
6. Validation of scientific knowledge
7. Objectivity and subjectivity in science.

The VOSE also included five concepts regarding the teaching of NOS, however, due to the current study population of high school students, these questions were not asked in this study. The pilot study involved open ended data based on VOSTS questions.

Following this administration, the author determined multiple issues with the VOSTS: some questions were too general, and the presence of ambiguous responses. The author also found that although respondents chose the same answer in response to questions, they did so for different reasons, resulting in a misconception regarding philosophical standpoint regarding NOS. Additional issues with the VOSTS were the presence of overlapping answers and that students were forced to choose one answer, which could have resulted in an incomplete picture of the respondents' views.

The second phase of VOSE development involved item development and testing. Each of the items was developed from the VOSTS, along with a Likert scale to facilitate scoring and the use of inferential statistics. Each of the questions was followed by a

series of statements for the respondents to rate. The last two questions of the survey involved two scientists with differing philosophical viewpoints. The pilot test for the initial draft was conducted with 120 biology students at a research university in Taiwan. From these results, the number of questions was reduced based on item difficulty and rate of responses to uncertain/no comment. When each respondent answered a question in the same manner, or when a large population answered the question with uncertain/no comment, the question was discarded. When a topic had less than three questions, the entire topic was removed as well. The author also investigated the possibility of order effect in the survey, and found that two questions were affected by order. Two panels of Taiwanese experts, each consisting of six professors who had published in the field of NOS examined the questions. The first panel was asked to establish face and content validity, while the second panel examined the researcher's interpretation for each item. When at least five of the six experts agreed with the researcher, the question was retained in the final version. This version was then presented to six college students and one high school student for clarity checking and the opportunity to express alternative positions regarding any of the items.

The third stage of the VOSE involved 302 junior and senior students across a variety of majors at two research universities in Taiwan. The same students were retested within 1-3 months for test-retest reliability, and they were interviewed following the test in order to justify positions and interpret the items. The interview questions were designed by the researcher and two graduate assistants conducted the interviews.

The VOSE is an empirically-derived survey, and thus was developed using a qualitative perspective that focuses on trustworthiness and authenticity of data. The

validity of the instrument came from the process used to develop the survey. The use of multiple data sources, including literature review, and the administration of VOSTS and VNOS contributed to the final instrument. Additionally, two expert panels of Taiwanese researchers in NOS ranked at Associate Professor or higher were used to establish content validity. Interviews with six college students and one high school student provided evidence of the clarity of the instrument, and in interviews, 83 of the 85 items had a 90% similar interpretation between respondent and researcher. Reliability for the instrument was also developed through the process of developing the instrument. The initial survey was open-ended, thus allowing respondents to provide a wide range of answers. As the survey was developed from the respondent's point of view rather than the researcher, the instrument was determined to be reliable. Additionally, Cronbach's alpha scores were established for each of the raw scores for each of the topics, and ranged from 0.34 to 0.80. The retest of the survey took place within one to three months after the initial administration of the survey, and yielded a score of 0.82 for test-retest reliability. The entire survey is presented as Appendix A.

Views on Nature of Science

The VNOS was developed by Lederman et al., (2002) in response to concerns regarding previously administered nature of science assessments. The instrument was designed to examine students' views on the nature of science, particularly the concepts most relevant to K-12 education: that scientific knowledge is tentative, empirical, theory-laden, and product of human inference, involves imagination and creativity, socially and culturally embedded, the distinction between observation and inference, lacks a universal scientific method, and the relationship between theories and laws. There are seven open

ended questions on the VNOS-B, which is the shortest of the three versions, requiring approximately 35-45 minutes for written completion. It should be noted that multiple concepts are contained within each question due to the interrelated nature of NOS themes. Within this study, the use of the VNOS-B was used as interview questions, with a graduate student or the researcher administering each of the interviews. Following the interviews, the primary researcher transcribed each of the audio recordings, and used a random number code to identify each student. Three graduate students, each with doctoral coursework in nature of science evaluated student responses for correspondence to contemporary views of the nature of science.

The issue of test validity has been raised (Aikenhead, Ryan, & Desautels, 1989; Lederman & O'Malley, 1990, from Lederman, et al., 2002) regarding nature of science tests as respondents and test developers often do not perceive test items in the same manner. Correspondingly, the instruments reflect the biases of the instrument developer, and may not reflect the true nature of science understanding of the respondent. The VNOS serves the purpose of examining respondents' views in depth through the free response format of the survey and the follow-up interviews that make up the protocol for the instrument. The interviews have the purpose of re-establishing validity to the instrument within the context of the study. The VNOS was administered to pre-service secondary science teachers (Abd-El-Khalick, 1998, from Lederman, et al., 2002) to determine the initial validity of the instrument. The researcher found that the respondents' interpretation of the questions matched the researchers during the follow up interviews. Construct validity was assessed (Bell, 1999, from Lederman et al., 2002) by gathering nine NOS experts from the fields of science education, history of science, and

research scientists and comparing their views on the VNOS-B with nine NOS novices with similar academic credentials in the fields of humanities and business. Bell found through the administration of the VNOS and follow-up interviews that the responses from the expert group were more representative of current NOS understandings at a rate of three times the responses of the NOS novice group. Thus, construct validity was established based on the instrument's ability to discriminate between groups of people.

The structure of the VNOS-B with interviews provided evidence for the validity and reliability of the instrument. The interview process allowed the researcher to compare oral interviews with written responses in order to address possible misunderstandings between researcher and respondent. Reliability for the instrument was primarily through inter-rater reliability, as multiple people examined the entire data set and reached consensus regarding each respondent's NOS views. The entire survey is presented in Appendix B. Table 2 shows the correlation between the VOSE and VNOS instruments.

Table 2

Congruence between VOSE and VNOS Constructs

Tenet of Nature of Science	Questions/statements on VOSE survey	Questions on VNOS-B
Tentative nature of knowledge	4. Even if the scientific investigations are carried out correctly, the theory proposed can still be disproved in the future.	1. After scientists have developed a theory (e.g. atomic theory) does the theory ever change? 2. What does an atom look like? How certain are scientists about the structure of the atom?
Empirical basis of scientific knowledge		2. What specific evidence do you think scientists use to determine what an atom looks like? 6. Is there a difference between scientific knowledge and opinion?
Creativity in science	3. When scientists are conducting scientific research, will they use their imagination? 5. Is scientific theory “discovered” or “invented” by scientists from the natural world. 6. Is scientific law ”discovered” or “invented” by scientists from the natural world?	4. How are science and art similar? How are they different? 5. Scientists perform experiments/investigations when trying to solve problems. Other than the planning and the design of these experiments/investigations, do scientists use their creativity and imagination during and after data collection?
Theory laden NOS	1. When two different theories arise to explain the same phenomenon, will scientists accept the two theories at the same time? 2 Scientific investigations are influenced by socio-cultural values. 8. scientists’ observations are influenced by personal beliefs, therefore, they may not make the same observations for the same experiment.	7. How are these different conclusions if all of these scientists are looking at the same experiments and data?
Relationship between theory and law	7. In comparison to laws, theories have less evidence to support them.	3. Is there a difference between a scientific theory and a scientific law? Give an example to illustrate your answer.

Using Science to Develop Reflective Judgment

The Reflective Judgment Model (King & Kitchener, 1994) is a stage theory that examines the changes in epistemic views that occur starting in adolescence and continue through adulthood. The Reflective Judgment Model (RJM) is based on two underlying observations: “1) individuals’ understanding of the nature, limits and certainty of knowing (their epistemic assumptions) affects how they defend their judgments; and 2) their epistemic assumptions change over time in a developmentally related fashion (King, 2008).” The development of students’ ability to make decisions and arguments in the face of incomplete or conflicting data should be a primary goal of science education, and is reflected in national reform documents. “Science distinguishes itself from other ways of knowing and from other bodies of knowledge through the use of empirical standards, logical arguments, and skepticism, as scientists strive for the best possible explanations about the natural world.” (NRC, 1996, p. 201) This benchmark regarding the nature of scientific knowledge captures the essence of reflective judgment, in which people progress through stages to develop the ability to gather and analyze data, and use data from various sources to make a reasoned argument.

There are seven distinct stages in the RJM, although these seven stages are grouped according to their level of reflective thought. The first three stages are pre-reflective thought, stages four and five comprise quasi-reflective stages, and stages six and seven constitute mature reflective thought. The *pre-reflective stages* are defined by a single, concrete truth. The pre-reflective stages place a great deal of faith in the “authorities” to determine the truth for us, and these people believe the “truth” that has (or will be) determined for them. As a result, people in these stages do not examine contrary

evidence, and in fact tend to believe the authority over the contradicting evidence. This view of knowledge hampered scientific development for centuries, as many people blindly followed the teachings of Aristotle instead of examining the world through their own eyes. Had they done so, they would have seen many examples of where the authority did not have the single, absolute truth.

The *quasi-reflective stages* are marked by an uncertainty in their belief system. They have advanced far enough to know that the authority cannot be right all the time, but they have not found a satisfactory paradigm to replace it. Although the role of authority is diminished, there is an understanding that authority is important (as biased in stage four, as experts in stage five), but they are no longer the sole source of information. The quasi-reflective person is rather cynical in stage four, as he or she believes that evidence is presented and molded to fit prior beliefs, rather than changing beliefs to fit the evidence. The person in this stage can examine the evidence for multiple viewpoints, but does not have the critical thinking skills needed in order to integrate the evidence into a cohesive knowledge base.

The *reflective stages* are marked by the shift from the passive receiver of information to the active producer of information. The authorities in this stage are seen as the experts in their field, and as such are the ones that construct solutions from the evidence produced. The person at this stage may often be one of the experts involved in constructing the knowledge, and he or she realizes that the knowledge changes according to the evidence presented. Thus, as the authority is the basis of all knowledge in the pre-reflective stages, the analysis of evidence becomes the basis of knowledge in the reflective stages. People in these stages realize that evidence and knowledge comes from

a variety of sources, and is able to interpret and analyze the evidence into a cohesive knowledge base. A summary table detailing the major characteristics, the role of authority, the role of evidence, views of knowledge and the concept of justification for knowledge in each stage is presented in Table 3 (Zeidler, et al., 2009, p. 93).

Table 3

Summary of Reflective Judgment Stages

Epistemic Cognition	Stages	Major characteristic	Role of authority	Role of evidence	View of knowledge	Concept of justification
Pre-reflective	Stage 1	Belief is concrete and single-category (there are no alternatives)	Authority and observation are the source of knowledge	Disconfirming evidence is denied. Belief does not depend on evidence.	Knowledge is absolute and concrete.	Beliefs do not need justification
	Stage 2	There is a true reality, but not everyone knows it.	Authorities know the truth, those who disagree are wrong.	Evidence is not needed to confirm belief, and cannot be used to disconfirm belief.	Knowledge is absolute, but not apparent to everyone at every time.	Justification by agreement with authority figure.
	Stage 3	Belief that authorities may not know the truth, but will someday.	Authority is the source of right answers, but there is no way to justify claims in areas of uncertainty	Evidence must be concrete and lead to a single answer.	Knowledge is certain in areas that are known, or temporarily uncertain in areas that are unknown	Right answers are provided by authority, other areas are unclear and defended by personal opinion.
Quasi-reflective	Stage 4	Understanding that one cannot know with certainty	Authority is often biased, they fit the evidence to their beliefs.	Evidence is used to confirm subject's prior beliefs	Knowledge is uncertain, there is always some ambiguity.	Justification provided by evidence that supports prior belief.
	Stage 5	Understanding that people cannot know directly, but can within a context based on subjective interpretation of evidence	Authorities are seen as experts in their field, perhaps limited by their perspective.	Evidence can be compared for different beliefs, but cannot integrate the evidence.	Knowledge is contextual because it is filtered through a person's perspective.	Beliefs are justified by evidence as it pertains to a particular context.
Reflective	Stage 6	Knowing is a process that requires action on part of the listener.	Authorities are involved in constructing solutions.	Plausibility of evidence and argument can be used to base beliefs for self.	Knowledge is based on information from a variety of sources.	Justification provided by comparing evidence and opinion, utility of solution.
	Stage 7	Interpretations of evidence and opinion can be synthesized into justifiable conjectures.	Subject is involved in constructing knowledge, and is aware that knowledge changes in light of new evidence.	Evidence provides logical solutions to problems, but may change in face of new or better evidence.	Knowledge is constructed by critical inquiry and evaluating evidence.	Beliefs are justified on the basis of probability, we can't know for sure, but wealth of evidence supports view

Prior research gathered on over 1,500 students from teenagers to middle adulthood show that high school students (n = 172) consistently showed pre-reflective thinking, with a mean score of 3.2, and that reflective thinking scores did not increase dramatically across the four years of high school education (King & Kitchener, 1994). The scores did tend to increase across the four years of college, and the highest reflective judgment scores were exhibited by graduate students. A reasonable question then becomes whether development in reflective judgment is a byproduct of having the appropriate educational experience, or merely age-related.

The RJM is a useful model for evaluating the efficacy of an SSI intervention due to the parallels between them. Both involve the use of ill-structured problems, which are problems which have multiple reasonable solutions. As such, both reflective judgment and SSI are based on the use of evidence to support decisions. Furthermore, there is an overlap of topics between the RJM and SSI. For example, one of the prompts for the Prototypic Reflective Judgment Interview deals with the safety of chemical additives. A dilemma for the Reasoning about Complex Issues Test involves the use of medications to treat depression. Both of these prompts are SSI in nature.

The implications for utilizing reflective judgment extend beyond the science classroom and into many areas, including “their personal lives, in the workplace, and in their communities...” (King and Kitchener, 1994, p. 54) With this in mind, King and Kitchener (2002, p. 55) provided some suggestions for including reflective judgment:

1. Show respect for students’ assumptions, regardless of the developmental stage(s) they exhibit. Their assumptions are genuine, sincere reflections of their ways of making meaning, and are steps in a developmental progression.

If students perceive disrespect or lack of emotional support, they may be less willing to engage in challenging discussions or to take the intellectual and personal risks needed for development.

2. Discuss controversial, ill-structured issues with students throughout their educational activities, and make available resources that show the factual basis and lines of reasoning for several perspectives.
3. Create many opportunities for students to analyze others' points of view for their evidentiary adequacy and to develop and defend their own points of view about controversial issues.
4. Teach students strategies for systematically gathering data, assessing the relevance of the data, evaluating data sources, and making interpretive judgments based on the available data.
5. Give students frequent feedback, and provide both cognitive and emotional support for their efforts.
6. Help students explicitly address issues of uncertainty in judgment-making and to examine their assumptions about knowledge and how it is gained.
7. Encourage students to practice their reasoning skills in many settings, from their other classes to their practicum sites, student organizations, residence hall councils, and elsewhere, to gain practice and confidence applying their thinking skills.

While the bulk of the research (King & Kitchener, 1994) on high school students' reflective judgment indicated that they consistently exhibit pre-reflective thinking ($n = 172$, 11 samples, 5 studies, $M = 3.2$), and that scores did not increase throughout high

school, research has provided evidence that high school students can develop more advanced stages of reflective thought following a year-long SSI curriculum in an anatomy and physiology course (Zeidler, et al., 2009). This study provided additional evidence regarding the efficacy of an SSI-based curriculum, and the study also added to the reflective judgment sample pool of high school students, which has traditionally focused on college students.

The reform movement in science of the last 20th century that has continued now into the 21st century brought with it a different vision of science education. Previously, science knowledge was meant to be learned, and the facts listed in the textbooks memorized. Laboratory exercises were primarily “cookbook” style, the students follow the prescribed steps in order, and get the expected results. More recently, however, the thought has been away from memorizing these disjointed facts and towards a more cohesive nature of science, which focuses on the larger context of science as socially constructed and socially influenced. Two separate papers (Dotger & Jones, 2007; Zeidler et al., 2009) raised the question of a link between reflective thought and NOS understanding, but neither one provides enough empirical evidence to draw a link between the two.

Assessing the Reflective Judgment Model

Reasoning about Complex Issues test

The RCI is a computer-based test developed by Patricia King and Karen Kitchener that involves five scenarios drawn from contemporary issues. The RCI is designed to measure reflective judgment from stages two to seven, which is the level of epistemological sophistication found in adolescents through adults. There are two main

sections of the RJI, the first gathers demographic and academic information, while the second section is the questionnaire. There are four general questions related to each situation: 1. an open ended question regarding their personal opinion on each issue, 2. a question asking either why experts disagree on the topic, or why the respondent believes the way he or she does, 3. statements derived from previous interviews on the topic answered using a five choice Likert scale. The five ranks include VS = very similar, S = similar, D = dissimilar, VD = very dissimilar, as well as M = meaningless, which is used as a reading check for respondents. The student responds to a series of statements based on how well each of the statements describes their own thinking on the subject, and 4. the student then rank-orders the statements from question three based on level of agreement with his or her own beliefs. Each of the questions was assessed by a computer program run by the authors of the test, who were responsible for analysis of the RCI data. The RCI has an internal consistency score of 0.61 for college freshmen and 0.57 for college seniors. The reflective judgment assessment was completed via a web-based application, so administration was based on the researcher's and teacher's ability to gain access to a computer lab for the administration of the RCI. A sample RCI protocol with directions is given in Appendix C.

Prototypic Reflective Judgment Interview

The PRJI is a semi-structured interview developed by King and Kitchener (1994) developed to assess respondents' level of reflective judgment on various scenarios. The PRJI involves three to five prompts, including the use of chemical additives, science and creationism as it relates to human origins, and the genetic determination of alcoholism (nature versus nurture). Each of the subjects is led through a series of seven questions

designed to elicit responses showing how students reason through these ill-structured problems, including the role of authority and the use of evidence, including conflicting data and claims. The seven standard questions are (King & Kitchener, 1994, p. 102):

1. What do you think about these statements?
2. How did you come to hold that point of view?
3. On what do you base that point of view?
4. Can you ever know for sure that your position on this issue is correct? How or why not?
5. When two people differ about matters such as this, is it the case that one opinion is right and one is wrong? If yes, what do you mean by “right”? If no, can you say that one opinion is in some way better than the other? What do you mean by “better”?
6. How is it possible that people have such different views about this subject?
7. How is it possible that experts in the field disagree about this subject?

The entire interview protocol with subject prompts are presented in Appendix D.

The traditional process involves the interview of subjects, the transcription of the interview, and coding the transcript to determine the stage of reflective thought. In all cases, the score is a three digit number. For a subject who is very consistent in one stage, the number may be repeated, such as 3-3-3. This would indicate that a subject consistently reasons at level three. More often, however, subjects reveal reflective thought indicative of two or more stages. In this case, the dominant stage is listed first, with secondary and possibly tertiary stages listed afterwards. An example would be 4-4-3, which would be indicative of a subject who typically reasons at level four, with some

instances of stage three reasoning. For data analysis purposes, each of the seven questions across three scenarios was rated by two independent scorers. These scores were used to determine instances of differential reasoning from the pretest to the posttest, rather than generating a holistic score for each student.

There are differences between the RCI and the PRJI that would necessitate the use of both instruments in this study. The first is the difference between recognition tasks and production tasks. The RCI is a recognition task, meaning that the possible answers are on the screen for the respondent, which consequently tends to result in slightly higher scores than the PRJI, which is a production task. The RCI is taken on the computer, with the results collected on a computer server, while the PRJI is an interview process. Therefore the RCI is more suited to gather large amounts of quantitative data, while the PRJI was given to a subsample of the respondents in order to gain more detailed epistemological profiles and richer qualitative data regarding reflective judgment.

Argumentation is Central to Science and Society

Describing Argumentation

The term argument has been used in many ways throughout the science education literature. The Oxford English Dictionary (from Driver, Newton, & Osborne, 2000) defines argument as “advancing a reason for or against a proposition or course of action (p. 291).” Jimenez-Aleixandre, Rodriguez, and Duschel (2000) identify three forms of argumentation: analytical, dialectical, and rhetorical. Analytical can be thought of as “formal logic” with an emphasis on inductive and deductive reasoning leading to a conclusion. Dialectical consists of “informal logic” as the reasoning leads to premises that are not immediately known. Rhetorical argumentation is oratorical in nature, based

on knowledge and persuasion. Zohar and Nemet (2002) define informal reasoning as, “reasoning about causes and consequences, and about advantages and disadvantages, or pros and cons, of particular propositions or decision alternatives. It underlies attitudes and opinions, involves ill-structured problems that have no definite solution, and often involves inductive (rather than deductive) reasoning problems (p. 38).”

There are three different views regarding the use of argumentation in science education: that argumentation is central practice of science (Newton, Driver, & Osborne, 1999), that argumentation is a mechanism for learning the epistemology of science (Sandoval & Millwood, 2008), and contextualized argumentation provides a mechanism for citizenship education (Zeidler & Sadler, 2008b). Much of the work in argumentation is based on the Toulmin (1958) Argument Pattern (TAP). The TAP focuses on the structure of arguments, rather than the content. Four main structures are data, claim, warrants, and backing. The data are the facts involved in making a claim, the claim is the conclusion whose merits are to be discussed, warrants are the reasons to connect the data to the claim, and backings are the basic assumptions that provide the grounds for the warrants. More complex structures in TAP include qualifiers and rebuttals. Qualifiers set the boundaries for the claim, while rebuttals are conditions under which the claim will not be upheld. Jimenez-Aleixandre, et al. (2000) examined the TAP and discussed the difference between field invariant structures and field dependent structures. TAP focuses on the field invariant structures, which are found across domains, while the field dependent structures involves what counts as each of the structures in TAP.

The role of argumentation is central to both scientific endeavors and society in general. Traditional conceptions of scientific knowledge formation have relied on the

processes of making observations and performing experiments, then communicating the results. These conceptions fall short of the mark, however, as they presume that the communication between the scientists who produce new knowledge and their readers tends to flow in one direction. In reality, there is typically much debate at the edges of “frontier science,” as the merits of new knowledge tend to be assessed from multiple perspectives. Newton et al., (1999) express this relationship when they argue, “Observation and experiment are not the bedrock upon science is built; rather they are handmaidens to the rational activity of constituting knowledge claims through argument. It is on the apparent strength of arguments that scientists judge competing knowledge claims and work out whether to accept or reject them (p. 555).”

Argumentation is also seen as a mechanism for learning the epistemology of science (Sandoval & Millwood, 2008). Both oral and written argumentation have been studied extensively, and the main goal of argumentation within a science education context is developing the ability in students to use evidence to support claims (Hodson, 2003, Kelly, Regev, & Prothero, 2008; Sandoval & Millwood, 2008; Yore, Florence, Pearson, & Weaver, 2006). Studies in oral argumentation (Driver et al., 2000; Erduran, Simon, & Osborne, 2004; Jimenez-Aleixandre, et al., 2000; Resnick, Salmon, Zeitz, Wathen, & Holowchak, 1993; Sadler, 2004) have mainly focused on group dynamics during collaborative inquiry or problem solving. The main finding from these studies was that students typically use claims without supporting evidence (Erduran et al., 2004; Jimenez-Aleixandre et al., 2000; Kelly et al., 2008; Resnick et al., 1993).

Argumentation has also been examined in student writing. Yore et al., (2006) reported that writing allows the opportunity to examine the structure of arguments, think

about ideas, and analyze the work of others. Within a science education context, written argumentation has the potential for students to understand scientific concepts through writing papers, reading other's papers, and reviewing other students' work (Bell & Linn, 2000; Kelly et al., 2008; Sandoval, 2003). As with oral argumentation, much of the focus on written argumentation has been on argument structure (Bell & Linn, 2000) and how students integrate data with text to formulate coherent arguments (Kelly et al., 2008).

If we are to educate the next generation of citizens, some of whom will become scientists, then it is our responsibility to provide them with the skills they need in order to be productive citizens, including argumentation skills. Some science educators believe that argumentation should take a central role in science education (Driver, Newton, & Osborne, 2000; Duschl & Osborne, 2002; Kuhn & Reiser, 2005; Newton, Driver, & Osborne, 1999; Osborne, Erduran, & Simon, 2004). Driver et al. (2000) believe there are many benefits to incorporating argument in the science classroom:

1. understanding the difference between observation and theory
2. understanding the epistemology of scientific knowledge
3. find out about science under consideration
4. distinguish between questions with a science basis and other types of knowledge
5. recognizing personal and social values that impact decision making
6. evaluating evidence from many perspectives.

From this list, it is apparent that utilizing argumentation as a strategy has the potential to realize many of the objectives listed in the national standards. Additionally, there is some evidence (Zohar & Nemet, 2002; Schwartz et al., 2004) that explicit argumentation may

enhance content knowledge, and others (Sadler & Donnelly, 2006) believe there is a reciprocal relationship between subject content and a context for learning science effectively. Zeidler and Sadler (2008b) argued that quality argumentation situations and the process of social discourse can promote conceptual understanding of the subject matter.

Additionally, there are many other areas of society (law, philosophy, and politics) that depend upon argumentation skills. Most notably, the existence of a participatory democratic society depends on the use of argumentation to foster the skills needed by students to become informed citizens. This idea was reflected in both national (AAAS, 1989) and international documents (OECD, 1998). The AAAS defines a scientifically literate person as

one who is aware that science, mathematics, and technology are interdependent human enterprises with strengths and limitations; understands key concepts and principles of science; is familiar with the natural world and recognizes both its diversity and unity; and uses scientific knowledge and scientific ways of thinking for individual and social purposes (1989, p. 4).

When combined with the OECD's definition of scientific literacy, these documents represent the conception that scientific literacy involves the use of scientific knowledge in order to foster decision making skills related to societal purposes. The use of argumentation in science education is critical to fostering this type of scientific outcome. Science educators (Davies, 2004; Hodson, 1993; Symington & Tytler, 2004) have also emphasized the link between democracy and scientific thinking. Symington and Tytler

characterize this thought with the idea that “education [should be] for ‘science in life’ rather than an education about science.” (p. 1403)

Additionally, Kuhn (1991) asked 160 people about three different topic areas: people who commit crimes while on parole, failure in school, and unemployment. The subjects were taken from a range of age groups from adolescence to elderly, with varying levels of education among the adults. She also included subjects she reasoned would provide domain expertise on the topic, including parole officers for the crime topic, teachers for the school failure topic, and philosophers (Ph.D. candidates in philosophy) with domain knowledge regarding general knowledge. She asked the subjects to describe and justify their theories (contemporary use of the term, as opposed to a scientific theory) and then probed those theories with alternative theories, counterarguments, and rebuttals. She found that 40% of the subjects could provide evidence to support their theories, 60% of the subjects could generate alternative theories, 50% of the subjects could comprehend evidence that falsified their position, and only 25% of the subjects could utilize an integrative rebuttal. The last two findings are particularly troublesome, as half of the subjects could not understand evidence that conflicted with their own belief and only a quarter of subjects could respond to an argument with an effective rebuttal. The ability to understand and evaluate multiple positions utilizing evidence should be considered fundamental in a democratic society, and when only a quarter of the subjects could engage in an integrative rebuttal, one could raise the question of the substance of discussions regarding this and other topics. This lack of argumentation awareness necessitates the need for explicit argumentation in secondary school. Zeidler and Sadler (2008b) argued that “educational programs and research focused on promoting

argumentation and character development should attend to how well students are able to articulate coherent and internally consistent arguments, recognize potential threats to positions and counter positions and form rebuttals.” (p. 212)

Argumentation Should be Explicitly Taught in Science

There is substantial evidence that argumentation is not taught in secondary science (Driver, Newton, & Osborne, 2000; Newton et al., 1999; Sandoval & Millwood, 2005). There have been many reasons given for the lack of argumentation in the science curriculum. Newton et al. (1999) interviewed fourteen experienced science teachers in England regarding the lack of argumentation in secondary science curriculum. The teachers offered a number of internal and external influences on their teaching priorities. Some of the internal issues were the classroom management skills needed to incorporate argumentation, the lack of quality materials for teachers, teachers’ skills and views of science, and the lack of teacher training in the area. Some of the external influences involved the time constraints necessitated by covering a national curriculum for external tests, the students’ and parents’ views of the need to fill “course books,” and students’ discomfort with participating in science discussions. Additionally, the methods used in the science classroom are not coordinated with the way scientific knowledge is constructed. “Finally, it is ironic that science, which presents itself as the epitome of rationality, so singularly fails to educate its students about the epistemic basis of belief, relying instead on authoritative modes of discourse.” (Scott, 1998 from Osborne, Erduran, & Simon, 2004) Ultimately, the incorporation of argumentation in the science classroom involves a shift in the classroom from what we know in science to how we know and why we believe (Duschl, 2008).

Despite the calls for argumentation to be incorporated in the science curriculum, and the realization that current science curricula do not typically make room for the implementation of argumentation, there needs to be a method for integrating argumentation skills seamlessly into the science curriculum. One possible solution is the incorporation of SSI to provide the context for science instruction. Some researchers have utilized SSI to foster argumentation skills over short time periods (one unit), however, there have not been any explicit attempts to investigate argumentation within the context of SSI over an extended period of time.

Summary

A majority of the science education research completed to date either provides a snapshot of students' abilities at a point in time or focuses on short term treatments to track developmental gains. Much of the research serves to "sound the alarm" and lament that students' abilities are not what we would like them to be. One could argue that it is unlikely that major cognitive shifts will occur over the period of a couple of weeks, and any short term gains may be lost over time. Few studies have utilized a treatment longer than a month to investigate scientific literacy. The present study shifts the conversation from a treatment of a single unit to a treatment that lasts a semester of high school (approximately five months). Further studies like this one will be needed to develop an explanation of student cognition as it develops throughout the course of secondary education, particularly within the context of novel curriculum, such as an SSI based course.

CHAPTER THREE: METHODS

Research Design

This study used a quasi-experimental design using students from intact high school classes randomly selected into a treatment (SSI curriculum) or comparison (traditional curriculum) group. Treatment classes were taught using a variety of SSI as the basis for learning content, while comparison classes were taught by the individual teacher in his or her normal routine. Both treatment and comparison classes included a laboratory component, as biology is a state-mandated laboratory science. The laboratory experience was consistent across treatment and comparison groups and determined by each of the classroom teachers.

Both qualitative and quantitative methods were used in order to gain a deeper understanding of the constructs under review. Historically, the use of mixed methodology helps to address two goals in research: the quantitative methodology to confirm a preexisting hypothesis, and the qualitative methodology to generate future areas of research and to explain the nature of the relationship between the variables (Teddlie & Tashakkori, 2003). In this study, the quantitative analyses derived from survey research predominantly served to describe trends in the classes as a whole, as well as serve the purpose of determining whether the SSI treatment enhanced the outcome measures in the study; the qualitative analyses derived from interview data provided evidence for nuanced changes among individual students and details on how the process takes place.

Purpose

This primary purpose of this study was to implement and analyze a semester-long (17 weeks including pre and post-testing) biology curriculum based on the use of SSI as the primary teaching method. There have been many studies that have examined the use of SSI over a short period of time (Walker & Zeidler, 2007; Zohar & Nemet, 2002), however, there are very few instances of the implementation of a semester long treatment. As the literature suggests that each of the outcome variables (nature of science understanding, reflective judgment, and argumentation proficiency) develops in stages over an extended period of time, differences within groups may be seen over the course of months that may not be apparent within the shorter time frame of a single unit.

Research Questions

RQ1. To what extent do students enrolled in high school science classes utilizing socioscientific issues show greater development in nature of science understanding compared to students in traditional high school science classes?

RQ2. To what extent do students enrolled in high school biology classes utilizing socioscientific issues show greater development in reflective judgment compared to students in traditional high school science classes?

RQ3. To what extent do students who are enrolled in high school biology classes utilizing socioscientific issues show greater argumentation skills compared to students in traditional high school science classes?

Population and Sample

The target population for this study included primarily ninth and tenth grade high school students enrolled in science classes. The sample was drawn from intact high

school classes of Biology I, typically offered in either the student's freshman or sophomore (grades 9 or 10) of high school. These classes came from a single school located in the Tampa Bay area of Florida. Two teachers from one high school volunteered to use their classrooms to implement the SSI curriculum to half of their classes, while maintaining their normal routine for the other classes in order to serve as a control. One of the biology teachers is a veteran with over fifteen years experience in the secondary school setting, and has worked with the researcher to incorporate debates in the past, while the second teacher has less than five years experience and has taken graduate level coursework in Nature of Science which involved theoretical background in socioscientific issues.

Based on demographic data from 2006-2007, the student population of the school in the study includes a minority rate of 26.80% (mostly African American), and a free and reduced lunch population of 25.58%. Each of the Biology classrooms serves the general population, and would have mainstream ESE students as well. As the Biology I course is one that virtually every student takes at the school, it could be reasonably inferred that the class population will mirror the school population for the freshman class.

The teachers were initially contacted about their participation in this study May 2008. Conversations continued over the course of the summer, as teaching assignments had the potential to influence this study. For the 2008-2009 school year, two of the teachers taught Biology I classes. One of the biology teachers had three classes of Honors Biology I, from which two classes were randomly selected to participate in the study, while the second teacher offered four sections of Biology I.

The students were selected on their basis of participation in one of these two teacher's classes. All students (who provided permissions) completed three quantitative measurements, a nature of science survey, a computer-based reflective judgment test, and a persuasive essay. Students were randomly selected from each class to participate in one of the three interview protocols. Depending on the number of students for the class who completed the IRB forms an individual student participated in either one or two interviews. The survey protocols examined aspects of the nature of science, reflective judgment, and argumentation. Having students interviewed for two different variables also allowed for an examination of the relationship between constructs more effectively. Approximately one hundred students turned in the IRB documents. Of these, thirty students were selected for pre-tests for each of the variables. Students who remained in the same classes were also interviewed at the end of the study. Due to schedule changes and student transfers, the students available for posttest interviews were fewer than the original thirty students who participated in the pretest interviews.

This study was approved by the University of South Florida's Division of Research Integrity and Compliance, and has also been approved by Pinellas County School's Department of Research and Accountability. Permission to access the sample was obtained by written parental consent as well as student assent. The researcher visited each classroom participating in the study in order to provide an overview of the study and distributed the parental consent and student assent forms. The classroom teachers collected the forms from the students during the class periods and returned them to the researcher. The researcher provided a phone number to address concerns from the parents. Approximately five phone calls were made regarding the study. The calls were

generally supportive and asked for more information regarding extra work for their son or daughter.

Variables

Argumentation

The process of using evidence and reasoning to support claims. The process of argumentation involves multiple people defending multiple viewpoints.

Nature of Science

A branch of the philosophy of science which deals with the characteristics of the scientific enterprise. The nature of science incorporates ideas regarding scientific epistemology, the scientific process, and the history and sociology of science.

Reflective Judgment

A model of cognitive development which describes changes in how people view and justify knowledge when faced with ill-structured problems.

Socioscientific Issues Curriculum

A curriculum developed in order to explicitly make the connections between science and society within the framework of contemporary social issues. Implicit within this framework is the expectation that argumentation and debate be included within the curriculum in order to develop moral and ethical reasoning.

Instruments/Measures

Each of the outcome variables under consideration were evaluated using mixed methods. Nature of science was examined through the use of a survey, the VOSE, and an interview protocol based on the VNOS-B survey. Reflective judgment was evaluated through the online survey, the RCI test, as well as an interview protocol, the PRJI.

Argumentation was evaluated through researcher developed rubrics for both written and interview protocols. The written/computer based instruments provided information about changes in the treatment and control groups as a whole, while the qualitative interviews provided information regarding individual's changes in thinking over the course of the semester. An organizer for the instruments is given in Table 4, while a detailed description of the instruments follows.

Table 4

Instruments used during the study

Construct	Quantitative	Qualitative
Nature of science	Views On Science and Education (VOSE)	Views on Nature of Science (VNOS-B)
Reflective judgment	Reasoning about Complex Issues Test (RCI)	Prototypic Reflective Judgment Interview (PRJI)
Argumentation	Written argumentation rubric (WAR)	Oral Argumentation Interview

Nature of science

Two separate protocols were used during the study. The VOSE, developed by Chen (2006b) provided a quantitative measure to seven aspects of the nature of science, while the VNOS-B (Lederman et al., 2002), usually a written protocol, was used to interview students during this study to gather qualitative data regarding students' views of the nature of science.

Views on Science and Education

Quantitative analysis was determined based on student responses on the VOSE survey, developed by Chen (2006a). The VOSE was designed to measure students' views regarding the nature of science, particularly: 1. tentativeness of scientific

knowledge, 2. nature of observation, 3. scientific methods, 4. hypotheses, laws, and theories, 5. the role of imagination, 6. validation of scientific knowledge, and 7. objectivity and subjectivity in science. There are also five questions regarding the teaching of science, which were not used in this study. The VOSE, as used in this study, consisted of ten question stems with multiple responses to each question stem. Each of the responses has a five point Likert scale, which ranged from strongly disagree to strongly agree. Scoring for the survey was based on agreement with contemporary views of the nature of science, with higher scores reflecting more contemporary views of nature of science, and lower scores reflecting more naïve views of the nature of science.

The VOSE was pilot tested in April of 2008 with ninety high school students to determine the readability of the survey. Students were instructed to complete the survey, as well as describe/underline areas which they do not understand. The results from this administration were reflected in the version of the VOSE given in this study. For the most part, the questions dealing with how science education is taught were removed, and words deemed difficult to student readers were defined. Four nature of science tenets were tested during the course of this study: tentativeness of scientific knowledge, nature of scientific knowledge, nature and comparison of theories and laws, and the use of imagination in science. A more detailed description of the VOSE questions used during this study follows. These tenets were chosen due to their general accessibility to high school students and their presence on both the VOSE survey and VNOS-B survey protocol.

Tentativeness of scientific knowledge. Contemporary NOS researchers (Osborne et al, 2003; Lederman et al, 2002; McComas et al, 1998; AAAS, 1989) have argued that

although scientific knowledge is durable it is still tentative in that many of our closely held beliefs are not absolute. The process by which new scientific knowledge may be formed can take on a revolutionary (Kuhn, 1962), an evolutionary (Popper, 1975/1998), or a cumulative form and the VOSE survey asks three questions regarding these stances (one question each). In each of these cases scoring is from one to five, with a score of five indicating agreement with the stance, and a score of one indicating disagreement with the given stance.

- 4. Even if the scientific investigations are carried out correctly, the theory proposed can still be disproved in the future.
 - A. Scientific research will face revolutionary change, and the old theory will be replaced.
 - B. Scientific advances cannot be made in a short time. It is through a cumulative process; therefore, the old theory is preserved.
 - C. With the accumulation of research data and information, the theory will evolve more accurately and completely, not being disproved.

Nature of scientific observations. The nature of scientific observations has traditionally been that observations are made by objective observers. However, the view that observers are bound by their knowledge of scientific theories and their biases has become the more predominant thought that scientific observations are theory laden. The VOSE survey includes five questions regarding the nature of scientific observations. Higher scores reflect a more contemporary view of the nature of scientific observations, that they are theory laden, while lower scores represent traditional thought that scientific observations are theory independent.

- 8. Scientists' observations are influenced by personal beliefs (e.g., personal experiences, presumptions); therefore, they may not make the same observations for the same experiment.
 - A. Observations will be different, because different beliefs lead to different expectations influencing the observation. (contemporary)

- B. Observations will be the same, because the scientists trained in the same field hold similar ideas. (contemporary)
- C. Observations will be the same, because through scientific training scientists can abandon personal values to conduct objective observations. (traditional)
- D. Observations will be the same, because observations are exactly what we see and nothing more. Facts are facts. Interpretations may be different from one person to another, but observations should be the same. (traditional)
- E. Observations will be the same. Although subjectivity cannot be completely avoided in observation, scientists use different methods to verify the results and improve objectivity. (contemporary)

Nature and comparison of theories and laws. The traditional thought regarding the development of theories and laws is they are discovered by scientists to describe a range of phenomena, while more contemporary thought reflects the idea that theories and laws are invented by scientists in order to make sense of a range of phenomena. The VOSE survey includes nine questions regarding the development of theories and laws, with high scores reflective of contemporary views of theories and laws, and lower scores reflecting a more traditional view of theories and laws.

- 5. Is scientific theory (e.g., natural selection, atomic theory) “discovered” or “invented” by scientists from the natural world?
 - A. Discovered, because the idea was there all the time to be uncovered. (traditional)
 - B. Discovered, because it is based on experimental **facts**. (traditional)
 - D. Invented, because a theory is an interpretation of experimental facts, and experimental facts are discovered by scientists. (contemporary)
 - E. Invented, because a theory is created or worked out by scientists. (contemporary)
 - F. Invented, because a theory can be disproved. (contemporary)
- 6. Is scientific law (e.g., gravitational law) “discovered” or “invented” by scientists from the natural world?
 - A. Discovered, because scientific laws are out there in nature, and scientists just have to find them. (traditional)

- B. Discovered, because scientific laws are based on **experimental facts**. (traditional)
- D. Invented, because scientists invent scientific laws to interpret discovered experimental facts. (contemporary)
- E. Invented, since there are no absolutes in nature, therefore, the law is invented by scientists. (contemporary)

One of the predominant misconceptions regarding theories and laws is the belief there is a hierarchy to them, with theories becoming laws when proven. This misconception is spread throughout many science courses, and tends to be a difficult misconception to correct. The VOSE survey includes four questions regarding a comparison between theories and laws.

- 7. In comparison to laws, theories have less evidence to support them.
 - A. Yes, theories are not as definite as laws. (traditional)
 - B. Yes, if a theory stands up to many tests it will eventually become a law, therefore, a law has more supporting evidence. (traditional)
 - C. Not quite, some theories have more supporting evidence than some laws. (contemporary)
 - D. No, theories and laws are different types of ideas. They cannot be compared. (contemporary)

Use of imagination in science. The idea that scientists use imagination during the scientific process is indicative of a contemporary view of the nature of science. This view includes the idea that scientists are creative not only during the development of appropriate questions and procedures for studying them, but also in the collection and analysis of data, as well as the interpretation of the data. The VOSE survey includes five questions regarding the use of imagination by scientists.

- 3. When scientists are conducting scientific research, will they use their imagination?
 - A. Yes, imagination is the main source of innovation. (contemporary)

- B. Yes, scientists use their imagination more or less in scientific research. (contemporary)
- C. No, imagination is not consistent with the logical principles of science. (traditional)
- D. No, imagination may become a means for a scientist to prove his point at all costs. (traditional)
- E. No, imagination lacks reliability. (traditional)

While the VOSE was used to generate statistical data to measure the sophistication of the nature of science, the qualitative portion of the nature of science assessment was completed using the VNOS-B survey as an interview format.

Views on Nature of Science

There are three different forms of the VNOS survey. Two forms are currently used, VNOS-B and VNOS-C. The VNOS-B is the shorter of the two forms, with seven questions on a range of nature of science principles. The VNOS-C contains ten questions on a wider range of NOS principles. For this study, due to the congruence between the NOS principles under investigation and the greater ease of administering the interview protocol, the VNOS-B was selected.

Tentativeness of scientific knowledge. Questions one and two investigated the tentative nature of science. Question one asked whether theories change directly while question two used an application question (the certainty of atomic structure) to investigate the tentative nature of scientific knowledge.

1. After scientists have developed a theory (e.g., atomic theory), does the theory ever change? If you believe that theories do change, explain why we bother to teach scientific theories. Defend your answer with examples.

2. What does an atom look like? How certain are scientists about the structure of the atom? What specific evidence do you think scientists used to determine what an atom looks like?

Nature of scientific observations. Questions six and seven examined the reasoning behind scientists' conclusions. The more sophisticated view of NOS is that scientists' views are a product of their experiences and biases, while the more naïve view is that scientists are completely objective.

6. Is there a difference between scientific knowledge and opinion? Give an example to illustrate your answer.

7. Some astronomers believe that the universe is expanding while others believe that it is shrinking; still others believe that the universe is in a static state without any expansion or shrinkage. How are these three different conclusions possible if all of these scientists are looking at the same experiments and data?

Nature and comparison of theories and laws. One of the most pervasive misconceptions is that a hierarchical relationship exists between theories and laws. Students often express the notion that theories are not proven, and that a proven theory becomes a law. This misconception is incorrect in two ways: belief that theories and laws serve the same purpose in science, and the second is that scientific knowledge could be considered "proven." Question number three on the VNOS-B examines students' views on the relationship between theories and laws.

3. Is there a difference between a scientific theory and a scientific law? Give an example to illustrate your answer.

Use of imagination in science. Students often believe that the scientific enterprise does not require the same level of creativity as the arts. However, from the framing of scientific questions to generating theories from data, the creative process permeates science. Questions four and five on the VNOS-B investigate students' views on the creativity of science.

4. How are science and art similar? How are they different?

5. Scientists perform experiments/investigations when trying to solve problems.

Other than the planning and design of these experiments/investigations, do scientists use their creativity and imagination during and after data collection?

Please explain your answer and provide examples if appropriate.

Reflective Judgment

There have been two major tools for measuring reflective judgment: the Prototypic Reflective Judgment Interview (PRJI), and the Reasoning about Complex Issues (RCI) Test. Both tools were used during the course of the study; the RCI provided quantitative data, while the PRJI provided qualitative data.

Argumentation

Argumentation was analyzed within two distinct situations, written and individual interview. The written argumentation was assessed by a researcher derived rubric (Written Argumentation Rubric, WAR) based on the previous work of Zohar and Nemet (2002) and Walker and Zeidler (2007). The interviews were examined by two graduate students familiar with the epistemology of argumentation and argumentation structure to examine changes in argumentation structure from pretest to posttest.

Written Argumentation

The WAR was developed using previous research from Zohar and Nemet (2002) and Walker and Zeidler (2007). The instrument was designed to measure argumentation skill based on specific criteria: number of justifications, the structure of argumentation, and subject matter knowledge. Zohar and Nemet (2002) provided the scoring for the structure of arguments. They used a range of scores for number of justifications using a range of 0-2 points (0 = no justification, 1 = one valid justification, 2 = two or more justifications) and 0-2 points for the structure of argument (0 = no argument presented, 1 = a simple argument or conclusion supported by at least one justification, 2 = complex argument with justification which is supported by another reason). Walker and Zeidler (2007) used a four point rubric to examine the use of subject matter knowledge in students' arguments (0 = no evidence or subject matter knowledge (SMK), 1 = incorrect evidence claims or SMK, 2 = non-specific evidence claims or SMK, and 3 = correct considerations of specific evidence claims or SMK).

The rubric was used to assess students' written argumentation skills as defined by the writing of a persuasive essay based on a scientific topic. High scores for each section of the rubric indicated a desired aspect of argumentation, such as more justifications, more complex argumentation structure, or correct use of subject matter when forming arguments. Lower scores indicated less desirable aspects of argumentation, such as less justification for beliefs, less complex argumentation structure, or less/incorrect subject matter used when forming arguments.

The teachers were responsible for assigning the written assignment, who then gave each of the essays to the researcher. Two science education doctoral students who

have had coursework in epistemology (including reflective judgment), socioscientific issues, and nature of science scored each essay, with the average score of the two raters indicating the final score for the student. Interrater reliability was initially assessed by scoring three essays on cystic fibrosis together and reaching a consensus on scoring. Following the cooperative scoring, each rater scored three essays individually. The initial inter-rater reliability was 75.3%, after which three additional essays were scored to achieve a rating of 93.3%. Each of the raters completed the scoring at the same time in the same location to facilitate discussion for problematic essays. The final score was determined by taking the average score of the two raters consistent with a scoring protocol from King and Kitchener (1994) for scoring the PRJI.

As the second posttest essay was of a different topic and scored at a separate time from the pretest essay, an interrater reliability score was determined for the fluoridation essay. One of the raters was involved in scoring the cystic fibrosis essays along with the researcher. Once again three essays were scored cooperatively with discussion regarding multiple points. A set of three essays were then scored individually, and the scores compared to determine an initial interrater reliability, which was 77.3%. Following further discussion, a second group of three essays was scored with an inter-rater reliability score of 95.0%.

Oral Argumentation

The oral interviews were conducted with the researcher and individual students to elicit responses to a structured interview protocol. Two graduate students rated each of the essays, particularly looking for differences between the pretest and posttest responses to the same questions related to desired aspects of argumentation, such as better position

and rationale, ability to take multiple perspectives, use rebuttals, or more use of subject matter when forming arguments. The researcher was responsible for administering the argumentation interviews. Following transcription of audio recordings, two graduate students scored each transcript, with the average score of the two raters indicating the final score for the student.

Data Collection

Teacher Selection and Training

The researcher made initial inquiries to potential teachers who might be willing to incorporate the SSI curriculum into their science classes. Discussion continued during the summer under the assumption that the teachers would be teaching Biology, as they had in the past. However, teaching assignments for some teachers changed, which excluded them from the present study. The researcher located two teachers who were willing to incorporate the SSI curriculum into their biology classes. One of the Biology I teachers has Honors level classes, while the other teacher has “regular” students. Biology I is primarily designed for younger (ninth and tenth grade) high school students, and students enrolled in these classes reflected a variety of intellectual development. Both teachers have bachelor’s degrees in science education, and both teachers have taken graduate level courses towards a Master’s degree in science education. If the SSI movement is to continue moving forward, the curriculum must be palatable to the vast majority of teachers who are not experts in the history, sociology, and philosophy of science education. The teaching schedule for the two teachers is listed in Table 5.

Table 5

Teacher Schedules for First Semester Biology

Teacher	Period 1	Period 2	Period 3	Period 4	Period 5	Period 6
1	---	Bio I (H) Comparison	Bio I (H) Not in study	---	---	Bio I (H) Treatment
2	Bio I Treatment	---	Biology I Comparison	Biology I Comparison	Biology I Treatment	

Each of the teachers involved in the study has had some exposure to SSI in the past. One of the teachers (teacher 1) has taken coursework as a graduate student in nature of science with a focus in socioscientific issues. The other teacher (teacher 2) has worked with the researcher on prior SSI units that formed the basis for national and regional presentations (Burek, Callahan, & Zeidler, 2004; Callahan, Robinson, and Fowler, 2006). In addition to the teachers' background knowledge, a training seminar presentation took place in August 2008 that provided the theoretical framework and goals of SSI in the classroom. The researcher initially interviewed each teacher individually regarding his or her beliefs about teaching science, specifically about the use of SSI and ethics within the science curriculum. This ten to fifteen minute interview focused on teachers' views on goals of science education, how science should be taught, what they knew about SSI in the classroom, and their current methods for teaching science. The initial training seminar lasted two hours and included three main topics: the theoretical framework behind SSI, general pedagogical concerns relative to incorporating debate/discussion in the science classroom, and a discussion of each of the cases to be used during the semester. The researcher discussed the theoretical framework behind a SSI curriculum, including the use of ethical issues, personal relevance of science content, and a dependence upon evidence-based reasoning for claims. The researcher and teachers

discussed the use of controversial and ill-structured problems as the basis for placing the content of the course within a context the students might know. Following the theoretical framework was a discussion regarding the pedagogy of implementing a SSI curriculum, including the use of a video from the National Center for Case Study Teaching in Science (2002) entitled *The Use of Case Studies and Group Discussion in Science Education*. This video presents the classroom management skills needed to conduct case studies in the science classroom. Many of the case studies prepared by the National Center for Case Study Teaching in Science can also be classified as socioscientific issues, due to their ill structured nature, their relevance to both society and science, and the inclusion of moral or ethical perspective taking. Both teachers had indicated that they had previously used discussion and debate in the classrooms, therefore the classroom management portion of the training was review to some degree. The discussion and video provided added nuances on the delivery of content and working within a discussion based classroom. Following the classroom management portion of the seminar, the curriculum for the course was discussed. It should be noted that each of the cases included both a teacher notes section as well as a student activity section. The use of the predesigned units of study provided some background to the teachers with how to facilitate the classroom. Each of the units in the course was reviewed with the teachers with suggestions for classroom practice. From the discussions during and following the training seminar, it was determined that each of the teachers had both the background knowledge and pedagogical strategies to incorporate the SSI curriculum into their classes.

Following this initial training, the researcher made weekly visits to each of the teacher's classrooms during class time to ensure that the SSI curriculum was being

followed appropriately, and had meetings with each of the teachers to suggest further techniques for furthering their pedagogical skills in this area. The researcher received a personal leave of absence from his full-time position as a teacher in order to fulfill the time commitments needed to complete the teacher and classroom observations necessary for study completion. The teachers were receptive to adjustments in strategy as the semester continued as well as to the use of SSI in general. One adjustment that was made by teacher two was the decreased emphasis on whole-group debate in favor of small-group discussions. The teacher reported that the students were willing to investigate the topics, but had difficulty debating in large groups. The use of smaller groups (4-5 students per group) with one student representing each interest group provided debates that were more successful.

There were further meetings prior to the implementation of each unit, in order to introduce the topic and discuss the fine points of the upcoming unit, as well as debrief the results of the prior unit. During this time, the teachers reflected on their teaching of the unit, and the researcher gave his perspective from classroom observations and respond to the teachers' reflections. Following the fifth SSI unit there was one final debriefing session to discuss the prior sixteen weeks as a whole. The researcher conducted individual interviews with each teacher regarding issues of teacher practices that helped the SSI treatment, what practices need further development, and student skills needed in order to complete the SSI assignments at a high level. The issue of student preparedness is outside the scope of the current study, however, the information gained from these discussions may lead to discussions between vertical teams of teachers in order to prepare students to participate in high-level argumentation scenarios.

Graduate Student Training

The researcher and one doctoral student in science education met prior to data collection to discuss the interview process and familiarize themselves with the interview areas. The doctoral student was familiar with the work of Lederman and the VNOS instrument from having taken coursework in the nature of science at the doctoral level. The researcher and the doctoral student discussed the differences between the typical administration of the VNOS-B (written with follow up interviews with a sample of the respondents) and the way the VNOS-B was used in this study (interview). The graduate student was given the protocol with introductory script as well as the VNOS-B questions. The graduate student conducted the majority of the nature of science interviews, while the primary researcher mainly conducted the reflective judgment and argumentation interviews. Throughout this process the researcher and the assistant discussed the interview process and made adjustments as necessary.

Prior to scoring the interviews, the researcher and two doctoral students (including the student who conducted the interviews) met to discuss the scoring of the argumentation essay and the interviews. Three student essays were selected and scored in a cooperative manner, discussing each of the criteria as we progressed. Following this initial group of three essays, the two graduate students scored three essays independently to achieve an individual inter-rater reliability. As the initial inter-rater reliability was below 90%, we discussed the scoring as a group and the two assistants scored an additional three essays independently. The secondary inter-rater reliability was greater than 90%, which was determined to be an acceptable level. The two assistants graded the essays at the same time and location and were able to discuss problematic essays.

The researcher and an additional two doctoral students met to discuss the scoring of the three interview protocols. Each of the assistants was familiar with the nature of science and reflective judgment protocols, as they had taken previous coursework in nature of science and epistemology. One of the doctoral students was the student who conducted the nature of science interviews, and the other doctoral student serves as the district supervisor for secondary science in a large county in Florida. His awareness of the state standards (which includes nature of science) and his coursework in nature of science and epistemology provided him the background needed to score the interviews. The researcher discussed the argumentation protocol with the assistants, which included the aspects of argumentation under examination. The graduate assistants were told to look for instances where a student had provided a substantively different answer from the pretest to the posttest. Each of the examples listed in the study were identified by at least two of the three scorers as a shift in response between the two tests

Curriculum Development

The development of the SSI curriculum took place prior to the implementation of the units in the fall semester. Each of the units has been chosen by the researcher with teacher input to gauge student interest in the topics. There are many sources for SSI units, including science education websites, previous course assignments, and creating new units. In this case, each of the units has been gathered from the National Center for Case Study Teaching in Science's website:

<http://ublib.buffalo.edu/libraries/projects/cases/ubcase.htm>. The Center, which is located at the State University of New York, Buffalo, grants permission for individual classroom teachers to print and distribute cases, and I asked for permission to ensure that my

situation falls within their guidelines. Each of these units was aligned to the state standards for Biology, thus the students learned the content through the multiple activities described in the unit. One should note that due to the complex nature of SSI, many of the units fit into more than one Sunshine State Standard, and often included interdisciplinary aspects as well, particularly with language arts, due to the high demands placed on reading and writing to complete the units. Each of the units were stored on compact disc and given to the teachers in advance of the school year, so that they had time to familiarize themselves with the curriculum. Each of the teachers was also given a teacher binder with teaching notes and student versions of the handouts.

Table 6

Correlation of SSI and Sunshine State Standards for Biology Curriculum

Sunshine State Standard	Nuclear power	Global warming	Stem cells	Transgenic crops	Pesticides
Nature of Matter	X				
1. All matter has observable, measurable properties.					
2. Basic principles of atomic theory					
Energy	X				
1. Energy may be changed in form with varying efficiency					
2. Interactions of matter and energy					
Force and Motion					
1. Types of motion may be described, measured, and predicted					
2. Types of force that act on an object and the effect of that force can be described, measured, and predicted.					
Processes that Shape the Earth		X			
1. Processes in the lithosphere, atmosphere, hydrosphere, and biosphere interact to shape the earth.					
2. The need for protection of the natural systems on Earth.					
Earth and Space					
1. The interaction and organization in the Solar System and the Universe and how this affects life on Earth.					
2. The vastness of the universe and the Earth's place in it.					
Processes of Life			X	X	
1. Patterns of structure and function in living things					
2. The process and importance of genetic diversity					
How Living Things Interact With Their Environment		X		X	X
1. The competitive, interdependent, cyclic nature of living things in the environment					
2. The consequences of using limited natural resources					
The Nature of Science	X	X	X	X	X
1. Scientific processes and habits of mind to solve problems					
2. Most natural events occur in comprehensible, consistent patterns					
3. Science, technology, and society are interwoven and interdependent					

There are some characteristics that define the pedagogy of a SSI curriculum, and effort was taken to ensure that these components were included within each of the modules. Such characteristics include the use of a moral or ethical component, the use of social discourse, and a resolution that results from the class discussion. The focus on a moral or ethical topic within the realm of scientific knowledge distinguishes the SSI movement from the previous Science-technology-society (STS) model. This use of an ethical issue tends to make the science personally relevant to the students, which is a primary goal of SSI instruction. The use of social discourse as an instructional tool allows students the opportunity to research a topic from multiple points of view, and then discuss the issue using their research and background science knowledge, as well as knowledge of economics, political science, religion, and sociology. This method presents the science as an integral part of society, rather than the traditional idea that science is separate from society. The social discussions lead to a class consensus on the issue, one that is socially determined and developed by the students. This aspect gave the students ownership of the knowledge being presented and the material being learned. The teacher's role was to serve as a judge or mentor to the students, rather than the dispenser of information. Each unit began with an introductory scenario that set the scene and made explicit connections between the content to be covered and the issue. A second feature of these units involved independent and small group research on the topic. Information literacy was not one of the constructs studied within the parameters of this study, however, the influence of technology is increasing, and students must be prepared to utilize a variety of applications and data sources in research. The third feature of each SSI unit was the use of small group discussion. An example would be a group of students given the same role in a

town-hall style debate. The participants needed to determine which points to emphasize, as well as which evidence was more convincing to support their position. The last characteristic assignment included a large group debate with other groups who held opposing points of view. This interaction with others has been a key feature in making SSI relevant to each student, as the knowledge is socially constructed.

Although each of these issues involved ill-structured problems within a controversial context, the assignments included within the units are varied. All of the units involved reading and analyzing text, involved the use of student research, and required the students to present information in small and large group settings. Each of the units also paid attention to the students' individual perspectives on each issue, which has been identified as an important aspect of utilizing SSI.

Although the focus of the study was the implementation of an SSI curriculum, there were many techniques the teachers used to present information. The implementation of the SSI curriculum comprised approximately fifty percent of the course time, with laboratory activities contributing another twenty five percent of the class time. The remaining twenty five percent of the class time included teacher lecture, discussions regarding content, tests, and multimedia presentations. As is the norm with curricula presented within the "real world" of secondary schools, the unexpected became routine. Fire drills, interruptions of class time due to announcements, activities, and PSATs all impacted the teachers and the implementation of the SSI curriculum, although each of the modules was presented in its entirety.

The SSI days were as varied as the normal school experience for the teachers. There were many activities involved in the delivery of an SSI curriculum, including

introducing the topic and assigning the students into groups, days where the students used computers and the internet to do research, debates and discussions, as well as conclusions and the expression of personal values and feelings regarding the issues. Although the debates are what people ordinarily think of when discussing SSI activities, the introduction and especially the research the students conducted prior to the debate played a large role in the success of the class and group discussions. The conclusion provided the students the opportunity to think about and express their personal views on the topics.

Biology I Curriculum

Below is a short description of each of the five units used during the Biology classes for the SSI treatment. An expanded description for each of the units is found in Appendix E, and these and other cases may be found at

<http://ublib.buffalo.edu/libraries/projects/cases/ubcase.htm>

Unit one: Tokaimura accident (Ryan, 2001).

This unit provides a chronological outline of a nuclear power accident at Tokaimura, Japan. The students follow multiple roles as they are led through the stages of the accident, from background material about nuclear power, through the conclusion and aftermath of the accident. Along the way the students make decisions regarding policy issues and discover more information about the process that goes into providing nuclear power.

Unit two: Re-enactment of the Kyoto Protocol. (Cowlshaw, Hunter, Coy, & Tessmer, 2006).

The debate behind whether the accumulation of greenhouse gases contributes to global warming has at times been contentious. The Kyoto Global Climate Conference

(1997) was an attempt to set world-wide guidelines reducing the amount of greenhouse gases produced by industrialized countries. The United States did not ratify the agreement, primarily due to the lack of restrictions on developing countries.

Unit three: Stem cells: Saving Superman (Rubin, 2003).

This unit regarding the use of stem cells in medical research revolves around the case of Christopher Reeve, the actor who fell from his horse and was subsequently paralyzed due to the breaking of two vertebrae in his neck. The fall caused massive trauma to his body, including paralysis from his shoulders down, the inability to breathe without a ventilator for months following surgery, and many other complications.

Although Reeve passed away in 2004, the controversy around the potential for stem cells to provide advances in areas such as spinal cord injuries, Parkinson's disease, diabetes, Alzheimer's disease, and many other disorders has not subsided. However, the most promising stem cell lines are fetal, and the possibility of aborted embryos providing stem cells draws the ire of the right to life contingent of Americans.

Unit four: Transgenic crops: Do you really know what you're eating? (Shew & Reese, 2007).

The issue of genetically modified foods has been explored (Walker & Zeidler, 2007) previously due to the links to many genetics concepts as well as having a potential for high interest for secondary students. This case study examines the fictitious story of a boy who suffers an allergic reaction to genetically modified corn present in taco shells and subsequently dies.

Unit Five: Pesticides: Can we do without them (Parendes & Burris, 2005)?

This unit places each person in the role of a county commissioner who must cast a deciding vote on whether to ban pesticides in Johnsonville County. The commissioner is provided information from a variety of stakeholders in the case: a Commissioner of Agriculture, the president of a homeowner's association, a representative from Natural Resource Conservation Service (NRCS), and the president from the local chapter of the Sierra Club.

Student Selection

Students were initially selected to participate in the study based on their enrollment in a participating teacher's Biology I class. The teachers gave the researcher a class roster for each class at the beginning of the semester. The researcher generated a random number code for each student. Following the return of parental consent and student assent forms to participate in the study, each student completed the written/computer based surveys in class. A random number generator was used to select three to four students from each class to participate in the nature of science interviews. The names of these students were provided to the graduate student who interviewed students as well as the teachers. A second random number generator was used to select three or four students from each class to participate in the reflective judgment interviews. The same process was followed for the argumentation interviews. In some cases the same student participated in two of the interview protocols, but none of the students participated in all three interviews.

A diverse group of students participated in the study. In addition to the variables measured (gender, race, ethnicity, and age) there was a variety of SES represented, as

well as mainstreamed ESE students in these classes. It should be noted in all tables that follow that these descriptions refer to students who completed the study from the beginning to its completion. Other students in these classes chose not to participate in the study, and there was immigration and emigration between the classes. These students are also not represented in the following tables. Table 7 describes the students based on gender. The treatment groups consisted of a total of twenty-two female students and twenty-seven males (44.9% female), while the comparison groups consisted of a total of twenty-seven females and twenty-four males (52.9% female).

Table 7

Description of Student Population: Gender

Period	Female (Percentage)	Male (Percentage)
1 (Treatment)	6 (37.5%)	10 (62.5%)
2 (Comparison)	6 (50.0%)	6 (50.0%)
3 (Comparison)	8 (57.1%)	6 (42.8%)
4 (Comparison)	13 (52.0)	12 (48.0%)
5 (Treatment)	4 (36.3%)	7 (63.6%)
6 (Treatment)	12 (54.5%)	10 (45.4%)

A variety of students based on race and ethnicity also participated in the study. In the comparison group 31.4% of the students identified themselves as members of a minority group (African-American, Asian-American or Pacific Islander, Native American, or Caucasian, while in the treatment group 14.3% of the students identified themselves as minorities. In the comparison group 15.7% of the students identified themselves as

Hispanic, while in the treatment group 12.2% of the students identified themselves as Hispanic. Table 8 provides the race and ethnicity data based on classes.

Table 8

Description of Student Population: Race and Ethnicity

Period	<i>Race</i>				<i>Ethnicity</i>	
	African-American	Asian-American	Caucasian	Native American	Hispanic	Non-Hispanic
1 (Treat)	0	0	16	0	3	13
2 (Comp)	3	0	9	0	0	12
3 (Comp)	2	1	11	0	3	11
4 (Comp)	8	1	15	1	5	20
5 (Treat)	0	0	11	0	1	10
6 (Treat)	2	4	15	1	2	20

Student age also had the potential to be a factor in the study. The average age of the forty-nine students in the treatment group was 15.0 years old, while the average age of the fifty-one students in the comparison group was 14.9 years old. However, two of the classes, periods one and three, had a class average over a year older than the other four classes. Table 9 represents the class average data for the study based on age.

Table 9

Description of Student Population: Age

Period (Group)	Average Age (in years)
Period 1 (Treatment)	15.9
Period 2 (Comparison)	14.7
Period 3 (Comparison)	15.9
Period 4 (Comparison)	14.6
Period 5 (Treatment)	14.6
Period 6 (Treatment)	14.6

Quantitative Procedures

Nature of science was assessed through the administration of the VOSE survey. The researcher discussed the structure of the VOSE survey with the teachers and answered any questions that arise. Teachers administered the survey in class. The teachers read the instructions to participants as the students followed along. The teacher was available to answer any questions that arise about the meaning of words, but did not provide other information to the students. Following administration of the VOSE survey, the teachers provided the surveys to the researcher, who coded each survey with the random number assigned to each student, and then used the number as the identifier for data analysis.

Reliability and validity data for the VOSE were obtained from Chen (2006b). She describes the survey as an empirically based questionnaire, developed from the learners' perspectives. Thus, Cronbach's alpha was not used as a main criterion for reliability, but

was used to evaluate individual questions for inclusion in the survey. The Cronbach's alpha scores for the questions ranged from 0.34 to 0.81. Validity was established through the use of two groups of experts and the researcher interviewing 24 subjects who participated in the survey. On 83 of the 85 items on the survey, over 90% of the respondents viewed the item consistently with the experts.

Reflective judgment was assessed through the administration of the RCI. The RCI was a computer-based survey which the researcher has received permission to use from the server administrator. The researcher was responsible for securing the computers needed to access the RCI website. The teachers and researcher were available to ask procedural questions regarding the survey, but did not answer any questions regarding its content. The server administrator provided a range of numbers as identifiers, which the researcher correlated to the student's random number identifier for the study. The internal consistency for the RCI has been reported in the low to mid 0.70s depending on the sample (King & Kitchener, 2004). The sample in this study produced lower Chronbach's alpha scores at 0.40 for the pretest, and 0.45 for the posttest. This may have been due to the use of high school students in the study, compared to college students who are the traditional population surveyed using the RCI.

Argumentation was assessed through the writing of an in class persuasive essay regarding a scientific issue. The students were given basic information about the topic from multiple perspectives (about two pages total reading), and the students was given the remainder of the period to complete the essay assignment. The teacher was responsible for administering the essay assignment. The teacher then provided the essays

to the researcher, who then added the random number code for each of the essays prior to data analysis.

Qualitative Procedures

Nature of science was assessed through the administration of the VNOS-B survey as an interview. A doctoral student who has completed coursework in the area of nature of science completed the majority of interviews with the students, which the remainder being conducted by the primary researcher. Each of the interviews was completed in an office workspace in close proximity to the science classrooms, and each of the interviews lasted between seven and fifteen minutes. The interviews were audio recorded using a digital audio recorder, and the student provided his or her name for the interviewer at the beginning of the interview. The audio recordings were then transcribed by the researcher, with the name of the student removed, and his or her random number inserted for identification purposes.

Reflective judgment was assessed through the administration of the PRJI, which has been used for many years as the standard for examining reflective judgment. The researcher conducted the majority of the interviews with the graduate student assisting as needed to complete the data collection. Each of the interviews took place in an office workspace located in close proximity to the science classrooms, with each interview lasting approximately between ten and twenty minutes. The procedures for collecting the data followed the protocol for the nature of science interviews.

Argumentation was assessed through the administration of an argumentation interview protocol. The primary researcher conducted the majority of the argumentation interviews, with the graduate student assisting as needed to complete data collection.

Each of the interviews lasted approximately seven to fifteen minutes and took place in an office space located in close proximity to the science classrooms. The researcher audio recorded the interviews, and transcribed the recordings utilizing the same protocol as the other two interviews. Table 10 provides information regarding the data collection throughout the study.

Table 10

Data Collection during the Study

Construct	Nature of Science	Reflective Judgment	Argumentation
Data Collection	Researcher and one graduate student	Researcher	Researcher and one graduate student
Quantitative Test	Views on Science and Education	Reasoning about Complex Issues	Written argumentation
Quantitative Sample	All students who consented to study	All students who consented to study	All students who consented to study
Qualitative Test	Views on Nature of Science (form B)	Prototypic Reflective Judgment Interview	Oral argumentation
Qualitative Sample	Five students from each class for pre-test, same remaining students for post-test.	Five students from each class for pre-test, same remaining students for post-test.	Five students from each class for pre-test, same remaining students for post-test.
Data Analysis	Two graduate students analyzed all VNOS transcripts	Two graduate students and researcher analyzed all PRJI transcripts.	Two graduate students and researcher analyzed written essays. Two graduate students analyzed transcripts from interviews.

Time Frame for Data Collection

The chronology for completing the survey is described in Table 11. Initially the teachers were identified in May of 2008 and trained in August of that same year. Informed consent protocols were distributed and collected during the first part of September through the middle of the month. Pre-test data were collected from the students following the return of the informed consent protocols in late September and early October. The administration of the three written instruments and the pre-test interviews took approximately six days with the researcher conducting interviews throughout the day and a graduate student helping as his schedule permitted. The nature of science interview (VNOS) took approximately ten minutes to administer. With the introduction and administrative details, approximately four students were interviewed within a fifty-minute class period. The reflective judgment (PRJI) was a longer administration (about twenty minutes) and approximately two were completed within a fifty-minute class period. The argumentation protocols also required approximately ten minutes and four interviews were completed within a class period. In total, ninety pre-test interviews were conducted (five students from each class for each of the three constructs). The administration of the curriculum followed the pretest data collection. Each unit required approximately three weeks with research time, laboratory activities, and basic content instruction incorporated into the lessons. The post-test data collection occurred following the completion of unit five. Due to the late start for the study, with the study receiving district approval in September and the consent period taking longer than expected, the study did not start until October and was completed in March of the following year. Within the time frame of the study the winter break occurred in

December as well as FCAT testing in February. The study was completed prior to spring break. The extended periods away from instruction may have played a role in the amount of material learned throughout the fifteen weeks of instruction. A number of students switched classes (periods and teachers) or withdrew from the school at the semester break, these students were removed from the study, hence the number of interviews for the post-test administration was fewer than the original set of thirty students interviewed for the pre-test.

Table 11

Timeline for Conducting Study

	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.
Teachers identified/ trained	X							
Informed consent protocol		X	X					
Pre-test data collection			X					
Science curriculum followed				X	X	X	X	
Post-test data collection								X

Procedures for Maintaining Confidentiality

Confidentiality of student data took the form of the use of random number identifiers for each student involved in the study. The student's names and district supplied student number (partial) were needed to gather raw data, however, the

researcher removed these potential identifiers before presenting information to the data analysts for review. The audio recordings were transferred from the audio recorder to the researcher's personal computer, and the transcripts of the interviews contained only the random number identifier.

For quantitative measures, the class data were reported for each outcome variable, while for qualitative measures, the random number identifier was used to report all data. The researcher is the only person who can match the number identifier with a student name. In no cases during the reporting of the data would an observer be able to determine which student was responsible for a particular piece of data.

Data Analysis

For each of the assessments involving quantitative analysis, including the VOSE, RCI, and both argumentation rubrics, descriptive statistics (means, standard deviations, and frequencies) were reported. Inferential statistics were conducted to determine both within groups and between group differences. As the class was used as the unit of analysis and the assumption of homogeneous variances was not met, the Wilcoxon tests were used to determine statistical significance. The Wilcoxon signed-rank test was used to calculate within group differences, while the Wilcoxon sum-rank test was used to calculate between group differences. Due to the small size of the sample groups (three classes for the control group and three classes for the treatment group), the alpha level was set at 0.10 in order to gain power (Stevens, 1999). Power is related to the possibility of making a Type II error. The raising of the alpha level beyond the typical 0.05 or 0.01 increases this power. Due to the lack of random assignment individuals to groups, there was no way to insure that the control and treatment groups were similar at the beginning

of the study. The use of pre-test data, however, provided some information regarding the equality of the groups before treatment. The SAS statistical software was used to complete all researcher-derived statistical analyses.

In addition to utilizing traditional inferential statistical analyses, this study also utilized the intra-sample statistical analysis (ISSA, Shaffer & Serlin, 2004). The ISSA deviates from traditional statistical analysis in that the requirement that the classroom be used as the unit of analysis is relaxed. In order to utilize this technique to describe the sample more fully, one loses the ability to generalize to an idealized population. For purposes of this study, the ability to investigate the sample under consideration was of greater importance than the power to generalize. The groups were examined for both between group differences on the pre and post tests, as well as within group differences on the pre and post tests. The Wilcoxon signed-rank test was used to measure within group differences, while the Wilcoxon sum-rank test (Mann-Whitney U) was used to measure between group differences.

Qualitative analyses were used to provide themes present in the student responses. It should be noted that these categories were predetermined in accordance with experts in the field. The categories established by Lederman et al. (2002) were used to examine student responses for the nature of science. King and Kitchener (1994, 2004) provided the framework for assessing reflective judgment, and argumentation draws upon the work of Zohar & Nemet (2002) as well as Sadler & Donnelley (2006) and Walker & Zeidler (2007). In each of these cases, pre-test and post-test interviews were conducted and transcribed with the respondents' pre- and post-test answers together to facilitate analysis. Doctoral science education students with coursework in nature of science and

epistemology analyzed each of the transcripts. Two doctoral students were assigned one of the three interview protocols, and the two students analyzed each of the interviews. The students were asked to look for instances in which the responses differed substantially from the pretest to the posttest in terms of sophistication. The reviewers were aware which interviews were from the pretest and which ones were from the posttest, but were blinded to comparison and treatment group. An example would be a reflective judgment interview in which a respondent provided a level two (pre-reflective) response to a question, and then provided a level three (quasi-reflective) response on the post-test. In all cases, sophistication refers to desired qualities as defined in this study, including a more sophisticated view of the nature of science, higher levels of reflective judgment, and more complex argumentation structures.

In this study techniques as outlined in Lincoln and Guba (1985) were used, such as semi-structured interviews and multiple analysts, to provide examples of student responses that changed in relation to the categories previously established in the literature for nature of science, reflective judgment, and argumentation.

Lincoln and Guba (1985) presented three terms that correlate to the traditional themes of reliability and validity. The first term is “credibility” Credibility can be built using a variety of methods, including activities that increase the likelihood that credible results will be determined. The first activity is prolonged engagement, which involves the researchers spending a sufficient amount of time in the classrooms in order to become unobtrusive. In this study prolonged engagement also had the benefit of checking on the pedagogy of the treatment as well, in order to provide evidence that the teachers were implementing the treatment protocol in the experimental classrooms, and not

implementing the treatment in their control classrooms. Triangulation enables the researcher to gather meaningful data from the study. Triangulation is the use of multiple sources, methods, and investigators. For this study, each of the constructs under evaluation was examined through two different sources. Nature of science was examined through the use of the VOSE (pen and paper) survey as well as follow up interviews (VNOS); reflective judgment was examined through the use of the RCI (computer based) and the PRJI (interview), and argumentation was examined through the use of essays (written) and interviews (oral). Each instrument that involved subjectivity in scoring involved the use of multiple investigators in the analysis of data, which for this study included the argumentation essay and each of the interview protocols. The VOSE survey was examined by multiple graduate assistants with coursework in nature of science to provide a consensus as to whether an “agree” response to each question stem resulted in a sophisticated or a naïve view of the contemporary nature of science. Three graduate assistants, including the researcher, independently analyzed the survey. Following the initial analysis, the researcher consulted the graduate students and consulted Chen’s (2006a) primary work to determine the categories and proper determination of the views.

The term “transferability” relates to the traditional term “external validity” as they both describe the degree to which results from a study can be applied to new situations. Although the nature of naturalistic research does not lend itself to making direct comparisons to other situations, there were procedures in place to increase transferability. The first was the use of purposeful sampling to maximize variation. Following the random assignment of students, the researcher determined a variety of students, based on ethnicity and gender, were included in order to maximize variation. The teachers also

reported a number of mainstreamed ESE students enrolled in the classes as well. As the Biology I class is taken by virtually all students in high school, the reasoning was that there would be a variety of students in the class. A variety of students participated in the study. The concept of “dependability” refers to the reliability of the study. Methods used to increase the dependability of this study included the triangulation described with credibility. An inquiry audit will be available in order to be open to evaluation from outside sources when questioned. These sources of information will include the raw paper and pencil instruments, computer generated printouts for RCI data, as well as both transcripts and audio cassettes for the student interviews. The collection of raw data will provide the independent observer the initial tools needed to determine the process taken in order to arrive at all knowledge claims.

Nature of Science

Data analysis for the VOSE survey took place based on the results of the descriptive statistics as described above. The VOSE data is ordinal, based on the ranks provided by a Likert scale response format. The pre-test data for each group were compared for similarity, while the post-test data reflects both differences in each group’s scores from the pre-test, and also differences between the groups.

Data analysis for the VNOS-B survey took place following a meeting by the graduate students in order to discuss conceptions of the nature of science, particularly what is meant by naïve and sophisticated views on nature of science as determined by Lederman, et al., (2002). Two raters analyzed each interview response. As this data was used to provide examples of potential sophistication in nature of science understanding, it

was determined that the two raters must agree on a score assignment to be used in the final data analysis for reporting.

Reflective Judgment

Raw data for the RCI were provided by the survey administrators, and the scores provided by the administrators were used to determine class and group data. Each student's score was imported to the researcher's statistical software in order to complete further data analysis, including inferential testing as previously reported.

Data analysis for the PRJI took place following a meeting by the graduate students to discuss the stages of reflective judgment, and to determine the process for determining a score for each question on the PRJI. Following the analysis of pilot data to determine inter-rater reliability, two scorers rated each student's transcript. Instances where the two raters agree that a respondent's answer became more sophisticated from the pretest to the posttest were used as qualitative examples for this study.

Argumentation

Data analysis for the written argumentation took place following a meeting by the scorers to discuss the rubric and to evaluate a small portion of the data set. Once the inter-rater reliability was above 90%, two raters scored each essay. The average of the two rater's scores was the final score given to the essay. In the instances that the scores deviated by more than one stage, a third rater (the researcher) scored the essay and all three scores were used in the determination of the student score.

Data analysis for the argumentation interview took place following a meeting by the scorers to discuss the interview protocol and the aspects of argumentation under consideration. Examples were given to the assistants describing each of the aspects, and

the two scorers then examined each of the transcripts focusing on changes from the pretest to the posttest, which was reported as changes in an individual's ability to construct arguments over the course of the semester.

Summary

This study used a mixed methodology consisting of a quantitative portion determined by scoring a survey (NOS) a computerized test (RJ) and an in-class essay (argumentation) to determine the effectiveness of a semester long Biology I curriculum based on the use of socioscientific issues to guide the curriculum. As the Biology I class is taken by virtually all ninth or tenth graders at the school used in the study, the course provided a diverse group of students engaging in SSI. Interviews conducted by the researcher and an additional graduate with individual students provided qualitative data that were used to determine growth on particular aspects of the outcome variables. Each of the quantitative instruments was analyzed using the Wilcoxon Signed Rank test to measure within group differences, and the Wilcoxon sum-rank test (Mann-Whitney U) was used to measure between group differences, run on SAS software. The use of this non-parametric test was judged appropriate given the small sample size (three comparison and three treatment classrooms) and the possibility that variance was not homogeneous across groups. Transcripts derived from the interviews were analyzed by two scorers in order to determine instances of more sophisticated explanations as well as identifying themes related to the outcome constructs.

CHAPTER FOUR: RESULTS

Introduction

Due to the mixed methodology of this study, the results section reports both data analysis and discussion of particular tables in chapter four, with the major themes of the study reserved for chapter five, as is traditionally done in qualitative research. As this study focuses on three distinct outcomes, each of the research questions are addressed and then answered in their original order of appearance in the study. The nature of science was investigated by using the VOSE survey to gather quantitative data and the VNOS-B as an interview to obtain qualitative data. Reflective judgment was studied by using two instruments, the RCI test to gather quantitative data and the PRJI to gather qualitative data. Argumentation was studied through the use of a persuasive essay on a science topic as well as an interview protocol to gather qualitative data. Both statistical data and interview data were utilized in order to provide more clearly the answer to each of the research questions. Interview data are represented with both a numeric value and the interview administration, with the numeric value representing the student's random identifier.

Research Questions

Research Question 1

To what extent do students enrolled in high school science classes utilizing socioscientific issues show greater development in nature of science understanding compared to students in traditional high school science classes?

This study attempted to answer whether the use of an SSI curriculum can produce changes in high school students' understanding of the nature of science. The VOSE survey examined multiple aspects of the nature of science including: tentativeness of scientific knowledge, the nature and use of theories and laws, the nature of observations, and the use of imagination in science. An average score (total) was derived from these four categories. Instances where the class average increased from pretest to post test are indicated in bold print. Table 12 shows the class means for these constructs.

Table 12

Pre and Post-test Mean VOSE Scores for Nature of Science Constructs (Post Test Data in Parentheses)

Class period (group)	Tentative	Nature of obs.	Imagination	Theory/Law	Total
1 (treatment)	3.22 (2.89)	3.01 (2.70)	3.07 (3.07)	3.24 (2.48)	3.14 (2.78)
2 (comparison)	3.31 (3.22)	2.90 (2.74)	2.76 (2.70)	2.79 (2.61)	2.95 (2.82)
3 (comparison)	3.67 (3.33)	2.54 (2.80)	2.90 (3.15)	2.66 (2.73)	2.94 (3.00)
4 (comparison)	3.18 (3.01)	2.86 (2.84)	3.00 (2.94)	2.78 (2.78)	2.96 (2.89)
5 (treatment)	3.33 (3.08)	2.65 (2.94)	2.99 (3.00)	2.67 (3.03)	2.92 (3.01)
6 (treatment)	3.46 (3.58)	2.53 (2.69)	2.92 (2.90)	2.53 (2.66)	2.86 (2.96)

From the descriptive statistical analysis, one can quickly discern that three groups appeared to make gains during the course of the study. Two of the groups, fifth and sixth periods, were part of the treatment group, while third period was part of the comparison group. Across all six groups, students scored highest on the tentativeness of scientific knowledge and the use of imagination in science on both the pretest (average 3.36 and

2.94 respectively) and the posttest (3.19 and 2.96 respectively). The NOS aspect regarding the use of theories and laws and comparison between theories and laws were combined into a total nature and comparison of theories and laws construct. Theory and laws were rated third lowest construct on the pretest (2.78) and decreased to the lowest construct on the posttest (2.72), while the nature of observations was the lowest rated construct on the pretest (2.74) and increased slightly (2.79) on the posttest.

The statistical data from the VOSE survey did not provide direct evidence to suggest that students make gains in nature of science understanding as a result of the SSI curriculum. Although the inferential statistics did not indicate a statistically significant result for the nature of science constructs, there were a few instances where student answers became more sophisticated from the pretest to the posttest. These instances are given below under each of the tenet headings along with the qualitative data.

Analysis of the individuals in the study through ISSA also did not indicate statistical difference. Median scores for the VOSE at the individual level for the pretest of the comparison and treatment groups were 2.95 and 2.85. The median scores for the posttests of the comparison group decreased to 2.90, while the score for the treatment group increased to 3.01. The score distributions did not differ significantly at the 0.10 level (2 tailed) between the groups either on the pretest (Mann-Whitney $U = 424.5$, $n_c = 39$, $n_t = 28$, $p = 0.1236$) or the posttest (Mann-Whitney $U = 640$, $n_c = 39$, $n_t = 28$, $p = 0.234$), although the change in the Z-score from -1.54 difference (treatment group scoring lower) on the pretest to +1.19 on the posttest (treatment group scoring higher) indicated that there was movement in a positive direction for the treatment group relative to the comparison group.

Within group differences were also examined at the individual level using the Wilcoxon Signed-Rank Test. The scores for the comparison group changed significantly at the 0.10 level (2-tailed) from the pretest to the posttest ($W = -257$, $n = 39$, $p = 0.0735$) in a negative direction. The scores for the treatment group did not change significantly from the pretest to the posttest ($W = 113$, $n = 28$, $p = 0.2005$).

The qualitative interviews were conducted prior to the study and again at the end of the study, about six months later. The general finding was that the students held a range of NOS views, and that these views did not change greatly over the course of the study. Below are student responses to questions on the VNOS, as they correlate to the constructs on the VOSE essay. Each of the responses includes the student number, followed by the question number on the VNOS, and the interview administration.

Tentativeness of Scientific Knowledge

Contemporary NOS researchers (Osborne et al, 2003; Lederman et al, 2002; McComas et al, 1998; AAAS, 1989) have argued that although scientific knowledge is durable it is still tentative in that many of our closely held beliefs are not absolute. The process by which new scientific knowledge may be formed can take on a revolutionary (Kuhn, 1962), an evolutionary (Popper, 1975/1998), or a cumulative form.

The VOSE results indicated that the tentativeness tenet of the nature of science was the tenet in which all students showed the most sophisticated reasoning, however for five of the six classes the posttest scores were lower than the pretest scores. The lone exception was period six, which increased from 3.46 to 3.58, or by 0.12 points. The VNOS interview provided examples in which students held naïve conceptions of the tentativeness of science throughout the study, students who held relatively sophisticated

views of the tentativeness of science throughout, and a couple of instances in which students made more sophisticated remarks at the end of the study. In these remarks, the “I” indicates the interviewer, while the “S” refers to the student.

I: After scientists have developed a theory (e.g. atomic theory), does the theory ever change (VNOS-B question 1)?

S: Yes the theory changes because with more evidence you can change theories.
(Student 10-1, post interview)

S: Um, I think if they really study it more, about whatever the theory is, um, that’s it’s possible for it to change, if they discover enough things. (Student 88-1, pre interview)

I: How certain are scientists about the structure of the atom (VNOS-B question 2)?

S: I believe they’re pretty certain about it because of the new microscopes they can see more and can see what it looks like more. (Student 10-2, post interview)

S: I would think they would be pretty certain on it since they, ah, research it, like, in depth and look through the microscope and actually can see in, and, yeah.
(Student 88-2, post interview)

These two students provided examples of naïve conceptions of the tentative nature of science. Although they were aware that scientific knowledge changes, they believed it does so through the accumulation of new evidence. Some of the students responded that scientific knowledge is tentative, and then expressed certainty with scientists’ current views of atomic theory. This showed the students’ inability to apply the basic concepts

of tentativeness to novel situations. Atomic theory is currently changing, as current atomic theory deals with quarks and leptons as the most basic subatomic particles, while current chemistry textbooks still tends to discuss atomic structure in terms of protons, neutrons, and electrons. Thus, the students were presented with a question regarding an issue that is currently changing, and has been changing over the past hundred years. The two following examples provided a more sophisticated view of the nature of science.

I: After scientists have developed a theory (e.g. atomic theory), does the theory ever change (VNOS-B question 1)?

S: Because we're always developing new technologies that can go deeper into what we've discovered so we teach what we know but it could change later.

(Student 23-1, post interview)

I: How certain are scientists about the structure of the atom (VNOS-B question 2)?

S: I believe they're pretty certain, um, because the way that they believe it can be used to prove how atoms react with one another. (Student 23-2, pre interview)

S: I'm pretty sure they are fairly certain but they can't be completely set.

I: What specific evidence do you think scientists used to determine what an atom looks like?

S: They've looked at – seen the basic structure with technology today and guessed what they can't see. (Student 23-2, post interview)

S: Um, I believe a theory does change, because as information is gathered over the years, there's new technology. The theory is changed and modified to go with that information. (Student 54-1, post interview)

I. How certain are scientists about the structure of the atom (VNOS-B question 2)?

S: I think they believe that there are tinier structures making up the protons and neutrons and electrons, but they haven't been able to identify them yet.

I: What specific evidence do you think scientists used to determine what an atom looks like?

S: Um, they've done a lot of research and experiments like we were talking earlier in the year about the electron, maybe, tunnel or something like that in London, or somewhere in that area, and they were shooting electrons at each other. Oh, it was the Electron Particle Accelerator, and they were trying to see if they could break the electrons into smaller pieces. (Student 54-2, post interview)

In this instance, student 23 provided a naïve explanation to the question regarding atomic theory during the pre interview, but provided a fairly sophisticated response during the post interview, including an understanding that the use of technology to gather new data or reexamine previous data occurs in science, as well as an understanding that inference is involved in theory generation. Student 54 provided a reasonable explanation during the post interview similar to the sophisticated response presented during the pre interview.

The students had a relatively sophisticated view of the tentativeness of scientific knowledge when asked directly, however the explanations then diverged into two major themes, 1. the traditional belief that new evidence arises to disprove old theories, and 2.

the more contemporary belief that technology is the driving force behind new scientific knowledge. The tentative nature of scientific knowledge was the VOSE construct with the highest average score across both the treatment and comparison classes, however this was not always seen in light of an application question during the interview. Most of the students indicated that scientific knowledge does change, particularly due to the invention of new technologies. Most students have a Popperian (1975/1998) view of scientific change, in that it happens evolutionarily – that changes happen over time, rather than an appreciation of the revolutionary concept of the tentativeness of scientific knowledge.

Nature of Scientific Observations

The nature of scientific observations has traditionally been that observations are made by objective spectators. However, the view that observers are bound by their knowledge of scientific theories and their biases has formed the more predominant thought that scientific observations are theory laden.

Three groups increased their VOSE scores on the nature of scientific observation tenet: third period (comparison) increased by 0.34 points from 2.54 to 2.80, fifth period (treatment) increased their score by 0.29 points from 2.65 to 2.94, and sixth period (treatment) increased their score by 0.16 from 2.53 to 2.69. The VNOS interview indicated that some students understood the sophisticated views of the nature of science in that scientists are bound to cultural and personal factors that guide their thinking.

I: How are these different conclusions if all of these scientists are looking at the same experiments and data (VNOS-B, question 7)?

S: Each scientist perceives a set of data as something different.... They're different people, they don't always believe in the same things. (Student 1-7, post interview)

S: Well, everyone portrays information in different ways, and one scientist may see one thing and use one piece of data to support their theory, where another scientist disregards that information and chooses another piece to emphasize in their theory.

(Student 54-7, post interview)

Both of these students emphasized the contemporary view of the nature of scientific observation in that scientists, while trying to be objective, are bound by their experiences, as well as personal, societal, and cultural factors. It should be noted that these two statements would likely be judged as quasi-reflective based on the reflective judgment model. There was an understanding that a difference of opinion existed, which is characteristic of both the quasi-reflective and reflective stages, however their responses that there was simply a difference of opinion (student 1) or bias on the part of the scientist (student 54) prevented these statements from being judged as reflective. A reflective response to this question would emphasize the probabilistic nature of the theories generated by scientists as a best interpretation of the data.

Nature and Comparison of Theories and Laws

The traditional thought regarding the development of theories and laws is that they are discovered by scientists to describe a range of phenomena, while more contemporary thought reflects the idea that theories and laws are invented by scientists in order to make sense of a range of phenomena. One of the predominant misconceptions regarding theories and laws is the belief there is a hierarchy to them, with theories becoming laws when proven. This misconception is spread throughout many science courses, and tends to be a difficult misconception to correct. Three groups (one comparison and two treatment) increased from pretest to posttest. Period three

(comparison) increased 0.07 points from 2.66 to 2.73, while two treatment groups (periods five and six) increased by 0.36 (from 2.67 to 3.03) and 0.13 (from 2.53 to 2.66) points respectively. The VNOS interviews provided many examples of misconceptions regarding the functions of theories and laws, as well as the relationship between them.

I: Is there a difference between a scientific theory and a scientific law? Give an example to illustrate your answer (VNOS-B, question 3).

S: Um, I think a scientific theory is something that changes over time, but a scientific law is something that never changes, because it's always true. Like Newton's laws, the gravitational law that will never change, we know. We can test that. That's going to be true every time. A scientific theory is just something that we think happened, like the theory of how the earth was made by the sonic boom [Big Bang Theory]. People think it's a fact by the sonic boom; other people think God made it. It's a theory. We won't really be able to find the right answer. (Student 10-3, pre interview)

S: Um, the difference between a scientific theory and a scientific law is a theory hasn't been around as long as a law, and a theory can be proven wrong with new information, but a law has been proven right every single time. And I think an example would be, um, Newton's laws of gravity and everything (Student 54-3, pre interview)

S: Scientific theory is just like, um, like, for example the Darwin's theory of evolution, and like him talking about the creature that walked over time. (Student 88-3, post interview)

S: Scientific law is proven by scientists and scientific theory is like an educated guess, pretty much.

I: Okay, and do you have any examples of those?

S: Um, scientific law would be Newton's Law.

R: Okay, and how about a theory?

S: Theory of the Big Bang. (Student 125-3, post interview)

The students continued the trend of low VOSE scores related to the epistemology and comparison between theories and laws with interview responses that highlighted the traditional view of theories and laws. Many students provided examples of theories that highlight society's contentious issues, such as evolution and Big Bang. Meanwhile, the same students provided examples of laws that are less contentious in society, such as the Law of Gravity. The examples these students used provided insight into what examples they are learning in science classes and hearing outside of school. As nature of science concepts are presented in the future, it might be useful to highlight theories that are more firmly established in society, as well discuss scientific laws that are newer and not as ingrained into the consciousness of students. This change may serve not only to change the epistemology of theories and laws, but with further explanation into these principles, may highlight the fundamental difference between the purpose of theories (explaining) and laws (summarizing).

Use of Imagination in Science

The idea that scientists use imagination during the scientific process is indicative of a contemporary view of the nature of science. This view includes the idea that scientists are creative not only during the development of appropriate questions and procedures for studying them, but also in the collection and analysis of data, as well as the interpretation of the data. Group three (comparison) increased on the VOSE by 0.25 points from 2.90 to 3.15 and group five (treatment) increased their scores by 0.01 from 2.99 to 3.00. The other four classes decreased from the pretest to the posttest on this tenet. The VNOS interview provided some insight as to how the students think about creativity in science.

I: How are science and art similar? How are they different (VNOS-B, question 4)?

S: I guess they're the same because you have to use pictures and drawings to illustrate what you've learned and, um, found out. (Student 1-4, pre interview)

S: Science is related to art – scientists use pictures and imagination to describe what they've discovered. (Student 1-4, post interview)

I: Scientists perform experiments/investigations when trying to solve problems. Other than the planning and the design of these experiments/investigations, do scientists use their creativity and imagination during and after data collection (VNOS-B, question 5)?

S: Well, if they collected all their data, and they have all their research, then I don't really see why they would need their imagination since they have the facts right in front of them. (Student 88-5, post interview)

S: I don't think so, but I think they should.

I: Okay, can you explain that a little bit more? Maybe provide some examples?

S: Um, kind of what it is, if you use creation and understand what you're doing, and think more about something, then you can expand what you're doing and research it more. (Student 124-5, pre interview)

Some students held to the traditional belief that there is no imagination in science, or were not able to identify instances where imagination would be used (Students 1 and 88).

The students who identified that scientists use imagination typically explained that this imagination occurred during the formation of hypotheses and procedures, and that the data analysis and theory generation required no imagination as the "facts were right in front of them (student 88)." One student (124) held that scientists did not use imagination, but that they should in order to "expand what you're doing and research it more." Thus this student understood the concept of creativity in science, even if the student did not believe it happens in real life.

Research Question 2

To what extent do students enrolled in high school biology classes utilizing socioscientific issues show greater development in reflective judgment compared to students in traditional high school science classes?

This study attempted to investigate whether high school students would attain higher levels of reflective judgment over the course of six months. As reflective judgment development tends to be an incremental process which increases more rapidly in college than high school, an underlying question is whether high school students are

cognitively unprepared to engage in higher levels of reflective thought, or is the traditional high school experience not conducive to advancing reflective judgment?

Reasoning about Complex Issues test

From the descriptive statistical analysis, the pretest data for the comparison and treatment groups were approximately equal (M = 4.27, SD = 0.17 and M = 4.20, SD = 0.51 respectively) while the treatment group appeared to increase on the post test (to M = 4.54, SD = 0.16) while the comparison group declined (to M = 4.15, SD = 0.31). Given the 0.46 point shift over the course of the six month treatment, further data analysis was warranted. Table 13 shows the overall means and standard deviations for both the pretest and the posttest on the RCI test for each period.

Table 13

Pre and Post-test Mean Scores for the RCI

Group	Pretest		Posttest		Change (rank)
	Mean	SD	Mean	SD	
Period 1 (treatment)	3.63	0.67	4.45	1.17	+ 0.82 (1)
Period 2 (control)	4.44	0.67	4.05	1.19	- 0.39 (6)
Period 3 (control)	4.27	0.85	4.50	1.01	+ 0.23 (2)
Period 4 (control)	4.10	0.23	3.90	0.89	- 0.20 (5)
Period 5 (treatment)	4.60	1.00	4.73	0.67	+ 0.13 (3)
Period 6 (treatment)	4.37	1.00	4.44	0.98	+ 0.07 (4)

The data were then analyzed using the Wilcoxon Signed Rank Test and found to not be statistically significant for within group differences for both control (n = 3, p = 0.25) and treatment (n = 3, p = 0.25) groups as well as between group differences (n = 3, p = 0.25). While the findings were not statistically significant, they did deviate from the accepted literature that suggests high school students average a 0.19 point increase (from 3.08 to

3.27 using the reflective judgment interview) over the four years of high school (King & Kitchener, 1994). It should be noted that the reflective judgment interview requires subjects to produce answers to ill-structured problems, while the RCI test requires subjects to recognize possible answers. Transitioning from the reflective judgment interview (RJI) to the RCI test typically produces an increase of one stage. Thus a student who scored a 3.25 on the RJI would likely score a 4.25 on the RCI. A 0.34 point increase (from 4.20 to 4.53) was measured over the course of six months using a socioscientific issues based curriculum. While further analysis must be done using larger sample sizes over longer term treatments to more accurately test for statistical significance, the evidence suggested that the SSI curriculum can be useful in providing high school students with a means to develop reflective judgment.

The data were also analyzed at the individual level using the ISSA methodology. Median scores for the comparison and treatment groups were 4.29 and 4.32 respectively, while the comparison group dropped to 3.87 and the treatment group increased to 4.54 on the posttest. The score distributions differed significantly at the 0.10 level (2 tailed) between the groups (comparison group was higher) on the pretest (Mann-Whitney $U = 234$, $n_c = 22$, $n_t = 24$, $p = 0.8729$) but not on the posttest (Mann-Whitney $U = 190$, $n_c = 22$, $n_t = 24$, $p = 0.2301$), although the change in the Z-score from +0.16 difference on the pretest to +1.2 on the posttest indicates that there was some movement in the scores in the positive direction for the treatment group.

Within group differences were also examined at the individual level using the Wilcoxon Signed-Rank Test. The scores for the comparison group did not change significantly at the 0.10 level (2-tailed) from the pretest to the posttest ($W = -47$, $n = 21$, p

= 0.4179). The scores for the treatment group also did not change significantly from the pretest to the posttest ($W = 75$, $n = 28$, $p = 0.2263$).

Prototypic Reflective Judgment Interview

In addition to the trends above, students were interviewed using the PRJI to elicit incremental advances in reflective judgment that might not be seen using the survey data. Most advances seen during the interviews (as would be expected) were from pre-reflective thought to quasi-reflective thought. Some examples are listed below with an explanation of the findings.

Religion and Science Issue

I: 4. Can you ever know for sure that your position on this issue is correct? How or why not?

S: Yes.

I: How?

S: Because it's been proven that we are a form of apes. (Student 139, pre interview)

I: 4. Can you ever know for sure that your position on this issue is correct? How or why not?

S: I don't know – maybe yes and maybe no. I'm not really sure.

I: And why would that be?

S: Because everybody – there's a lot of scientists that study evolution and some of them say we evolved from animals and others say we didn't. (Student 139, post interview)

The student progressed from a pre-reflective stage that is marked by certainty of knowledge to a low quasi-reflective stage where knowledge is uncertain and marked by ambiguity. In this instance, the student responded in the pre-test interview that it has “been proven that we are a form of apes” and during the post-test interview responded that there are scientists on both sides of the issue. This shift from a certainty of knowledge to an understanding that knowledge is not concrete is a major distinction between pre-reflective and quasi-reflective thinking. It has been noted previously (Zeidler et al., 2009) that the position a respondent takes does not factor into evaluating a statement as reflective or not; rather it is the explanation a respondent uses to justify his or her position that determines the coding of a response.

Alcoholism issue

I: 5. When two people differ about matters such as this, is it the case that one opinion is right and one is wrong?

S: If they’ve done some research then one could be right and one could be wrong, but with more research then one is probably the right answer.

I: What do you mean by “right”?

S: They can prove it or they can change the way those people live and are not alcoholics any more then that one’s right (Student 125, Alcoholism, pre interview)

I: 5. When two people differ about matters such as this, is it the case that one opinion is right and one is wrong?

S: No it’s just that one is more reasonable and makes more sense.

I: And what would make one opinion more reasonable or more sense?

S: If you compare people's lives and see who actually became an alcoholic and you would probably see that the experiences are probably bad compared to someone who is not an alcoholic. (Student 125, Alcoholism, post interview)

During the course of the pre-test interview, the student noted that through research the right answer could be found, which indicates pre-reflective thinking (stage 3). During the course of the post-test interview, the student shifted from an absolute view of knowledge to one that was based on probability and reason. The student changed her view of testing from providing a correct answer to a view that testing provides evidence for researchers to utilize in the process of drawing conclusions. She, in this instance, provided a high quasi-reflective (stage 5) answer, which was substantially higher than the stage 3 answer she provided before the treatment.

Research Question 3

To what extent do students who are enrolled in high school biology classes utilizing socioscientific issues show greater argumentation skills compared to students in traditional high school science classes?

This study attempted to answer whether the use of a socioscientific issues based curriculum could have positive effects on argumentation skills than what is traditionally seen at the high school level. There were three criteria examined for the written argumentation exercise: the use of justifications, argumentation structure, and the use of subject matter knowledge. Both justifications and structure were scored from zero to two, with two being the highest score (Zohar & Nemet, 2002). Subject matter knowledge was scored from zero to three, with three being the highest score (Walker & Zeidler, 2007).

The oral argumentation interview was based on number of justifications, position and rationale, and subject matter knowledge. The findings for each of these exercises are presented below.

Written Argumentation

Judging by the responses on the initial posttest, it became clear that the students were not motivated to perform well. This was indicated by a two to three point drop in the overall average for each class. The teachers indicated that many of the students were not motivated to complete the cystic fibrosis essay for the posttest. The teachers reported that many students did not want to write responses to an essay prompt they had already been given earlier in the year (pretest), and/or lack of motivation on the students' part following standardized testing and prior to Spring Break. After consulting with the teachers and major professor, it was decided that an alternative essay assignment (on fluoridation) might be beneficial in eliciting actual abilities on the argumentation essay. For the second posttest administration, four of the groups increased (two treatment and two comparison), one group remained the same (comparison), and one group decreased (treatment). Table 14 shows the pretest and posttest descriptive data for each of the classes.

Table 14

Pre and Post-test Mean Argumentation Essay Scores

	Pre-test		Post-test (CF)		Post-test (Fluoridation)			
	Mean	SD	Mean	SD	Mean	SD	change (Rank)	
Period 1 (treatment)	4.25	1.06	2.00	2.82	4.75	1.44	+0.50	(3)
Period 2 (comparison)	5.44	0.98	4.73	1.29	5.44	1.44	0	(5)
Period 3 (comparison)	5.17	0.76	3.00	2.65	5.50	0.55	+0.33	(4)
Period 4 (comparison)	4.00	1.11	2.89	1.49	5.30	1.26	+1.30	(2)
Period 5 (treatment)	5.00	1.68	3.00	2.65	6.33	1.21	+1.33	(1)
Period 6 (treatment)	5.79	0.98	5.34	1.13	5.32	1.10	-0.47	(6)

The changes in the scores from the pretest to the posttest could have been a factor of utilizing a different prompt than the pretest, and the lack of a trend between which groups showed greater increases indicated a lack of statistical significance for this construct. The treatment groups had the highest change, the third highest change, and the lowest change from pretest to posttest.

The data were also analyzed at the individual level using the ISSA methodology. Median scores for the comparison and treatment groups were 5.0 and 6.0 respectively, while the comparison group increased to 5.5 and the treatment group decreased to 5.5 on the posttest. The score distributions differed significantly at the 0.10 level (2 tailed) between the groups on the pretest (Mann-Whitney $U = 501$, $n_c = 31$, $n_t = 26$, $p = 0.0629$) and the posttest (Mann-Whitney $U = 412.5$, $n_c = 31$, $n_t = 26$, $p = 0.6892$).

Within group differences were also examined at the individual level using the Wilcoxon Signed-Rank Test. The scores for the comparison group did not change significantly at the 0.10 level (2-tailed) from the pretest to the posttest ($W = 83$, $n = 31$, p

= 0.18). The scores for the treatment group also did not change significantly from the pretest to the posttest ($W = -37$, $n = 25$, $p = 0.60$).

Oral Argumentation

The argumentation interviews examined student responses to two different scenarios: cystic fibrosis (CF) and cheating on tests. Four criteria were examined: position and rationale, multiple perspective taking, the ability to utilize rebuttal, and subject matter knowledge. The first three constructs were scored from zero to two, with two being the highest score, and subject matter knowledge was scored from zero to three, with three being the highest score. Very few of the students reached a maximum score on any of the constructs. The maximum position and rationale score correlated to an extended argument that includes a claim with supporting grounds (reason and evidence). All students were able to construct a reason for their belief, but only one student provided the supporting evidence to back up their claim.

After seeing the disease and seeing how many problems it causes, physically and medically. I know someone with cystic fibrosis and it's really hard for them to live like – they have tons of hospital runs and they go to school normally, which I don't know how, but people make fun of them. A lot of people help them, but like life is really hard and you can't participate in many activities like normal kids – and they rarely live past forty, even if they reach forty. So I think it should be aborted because rather than living a life of twenty or thirty years, in pain and suffering – it wouldn't even be happy, so I think they should (unintelligible).

(Student 3, CF Issue, pre interview)

This student discussed the physical and medical problems associated with cystic fibrosis, and then provided evidence of the different problems: hospital trips, difficulties in school, and the inability to participate in activities. The student also discussed the short life span of the cystic fibrosis patient as her claim that a fetus that may have cystic fibrosis should be aborted. Many other students provided a short reason without extensive grounds to support their position.

I: Do you think they should perform an abortion?

S: Yeah, if they know it's going to have CF because it's kind of mean to have the baby suffer and not even live that long. (Student 56, CF Issue, pre interview)

I: Do you think they should perform an abortion?

S: Yes because they should not – I don't think they should make it a burden onto their child and onto the kids – their grandchildren and stuff like that.

(Student 145, CF Issue, post interview)

I: Do you think they should perform an abortion?

S: No.

I: Why not?

S: Because it's as if taking another person's life away. (Student 145, CF Issue, pre interview)

I: Do you think they should perform an abortion?

S: No because even though it has the disease it should still have a chance to live.

(Student 185, CF Issue, post interview)

For the CF Issue, the majority of the positions debated the issue of quality of life against the morality of abortion. While few of the positions included an extended argument, the students were able to define a position either in favor of the parents keeping the baby or against having the baby suffer over the course of its lifetime.

The pattern of students providing claims without grounds continued for the cheating dilemma. Students were articulate a reason for their belief, but did not support that reason with evidence.

I: Do you think Rick should tell the teacher that Andy cheated? Offer reasons for your position!

S: Yeah, because if he keeps on doing it then he'll get in more trouble. (Student 81, Cheating Issue, pre interview)

I: Do you think Rick should tell the teacher that Andy cheated? Offer reasons for your position!

S: I don't think it's right for Rick to go and rat on Andy. I think Rick should try his best to convince Andy to confess, but if at the end of the day Andy just won't then I think it's just best if Rick just leaves it alone. (Student 206, Cheating Issue, post interview)

Summary of Results

Although the study did not produce any instances of statistical significance, some students showed more sophisticated reasoning from the pretest to the posttest measures.

The students were relatively energetic and optimistic at the beginning of the study, but during the spring semester their enthusiasm waned and for some students, their motivation to do well in class subsided. This study was not completed during the first semester as planned. Some factors for this were the length of time required to complete the informed consent process and the length of time needed to finish the SSI units. The treatment therefore, continued until March 2009, right before the state standardized tests, and the posttest interviews were conducted between the standardized tests and Spring Break periods, which was not conducive to achieving maximum results.

The nature of science was tested through two protocols, the administration of a survey (VOSE) and an interview using the VNOS-B as a guide. The VOSE data indicated that most students were approximately halfway between traditional and contemporary nature of science views, and these views did not change greatly over the course of the study. Three groups improved their scores on three out of four tenets of NOS as tested by the VOSE. Period three (comparison) performed better on the posttest regarding the nature of scientific observations, the use of imagination in science, and the nature and comparison of theories and laws. Periods five and six (treatment) both improved from pretest to posttest on the nature of scientific observations and nature and comparison of theories and laws tenets, but period five also improved on the use of imagination in science tenet, while period six improved on the tentativeness of scientific knowledge. Overall, two of the treatment groups (periods five and six) made the largest gains on the overall VOSE survey, with the comparison group (period three) making smaller gains on the VOSE. Three groups did not improve on any of the constructs, including period one (treatment), and periods two and four (comparison). Period one

(treatment) had the largest decline of the three groups. The improvement made by period three and the lack of improvement seen in period one were enough to show a lack of statistical significance when using the nonparametric Wilcoxon Signed Rank Test. The VNOS interviews revealed that many students had more naive conceptions of the nature of science than indicated by the VOSE survey. Part of this discrepancy may be due to the recognition task nature of the VOSE survey, while the VNOS was a production task.

Reflective judgment was also tested using two different instruments: the RCI computer based test as well as the PRJI protocol. Overall, the comparison group decreased from the pretest to posttest, with periods two and four showing declines (0.39 and 0.20 respectively), although period three showed an increase of 0.23 points. Overall, the treatment group increased slightly over the course of six months, with all three periods posting gains of 0.82, 0.13 and 0.07 respectively. Although the results were not statistically significant, this may have been an artifact of the study (small number of classes) as the treatment group showed greater increases than would be expected in the literature. One comparison group showed gains larger than two of the treatment groups, which may have led to the lack of statistical significance. As the students taking the RCI included two honors classes and four regular classes with a fair amount of mainstream ESE and low level readers, it could be reasonably inferred that this sample is representative of the high school population as a whole. The PRJI reflected the results from the RCI, with small increases made across some questions, mainly from pre-reflective thought to quasi-reflective thought.

Argumentation was also tested through the use of two protocols, an argumentation essay as well as an argumentation interview protocol. The essay assignment suffered

initially from being the same protocol as the pre-test. Students were not motivated to complete the essay assignment having completed the assignment earlier in the year. This obstacle was alleviated by assigning the students a second, different essay which the teachers reported would be part of their grade. The responses from the second essay more accurately reflect a maximum effort given by the students on the assignment. Four of the classes (two comparison and two treatment) increased from pretest to posttest, one comparison class (period two) scored the same, and one treatment class (period six) decreased from pretest to posttest. The class that decreased had the highest score on the pretest essay, but only third highest score on the posttest. Although the scores did not increase greatly, it should be noted that the students had not engaged in any formal debate or argumentation prior to the study. The interviews reflected a number of students who were able to construct simple arguments consisting of a reason to back up a claim. Many students were not able to include evidence for their argument, and merely restated their own argument when faced with a counter position.

The SSI treatment provided some instances in which advances in NOS, reflective judgment, and argumentation were made, however, the advances were not as widespread as one would hope over a long term treatment. Therefore, one could argue that the length of the treatment is not the deciding factor in student achievement across these constructs. Explicit instruction for each of these constructs within the context of a socioscientific issues curriculum may provide gains that a socioscientific issues curriculum alone does not seem to provide.

CHAPTER FIVE: DISCUSSION

Introduction

In the preceding chapter, the presentation and analysis of data were conducted. Chapter five consists of a discussion of the findings, implications for educational practice, recommendations for further research, and conclusions. The purpose of these sections are to expand the analysis from chapter four, forge direct links between research and practice, as well as provide additional directions for future research studies. Finally, a concluding statement describes the scope of the present study.

Discussion of the Findings

This study attempted to close some of the gaps in the scientific literacy literature by providing a context from which students were exposed to a semester long treatment. Previous literature (NCES, 2004; OECD, 2006) suggested that students from the United States were underperforming in terms of scientific literacy. Additionally, fundamental literacy has been seen as central to scientific literacy (Norris & Phillips, 2003), and others (Zeidler, 2007) have argued that moral reasoning and character education should be included within the definition of scientific literacy.

The idea of scientific literacy within the context of secondary education is part of the education of a knowledgeable populace ready to make decisions on issue of scientific import in contemporary society. Kuhn (1993) wrote that all students must be capable of scientific thinking and understand the importance of science in their everyday lives, while others (Davies, 2004; Hodson, 2003; and Symington & Tytler, 2004) have argued that

students' epistemological growth is necessary for the continuation of a democratic society.

The SSI movement focuses on the incorporation of science issues with social relevance. One of the goals of this study was to successfully implement a semester long Biology I curriculum with SSI as the focus for learning biological concepts, and this goal was accomplished. Students remarked that they appreciated the opportunity to express opinions, that the units provided an aspect of personal relevance, and there was an increased interest in the subject matter (Callahan, 2009). Interviews with the teachers echoed these sentiments. Previous researchers have investigated the benefits of a socioscientific issues based curriculum (Fowler et al., 2009; Zeidler et al., 2006; Zeidler et al., 2009). The primary goal of this study was to investigate the relationship between socioscientific issues and three outcomes related to scientific literacy. This section discusses the implications of utilizing a socioscientific issues based curriculum.

1. To what extent do students enrolled in high school science classes utilizing socioscientific issues show greater development in nature of science understanding compared to students in traditional high school science classes?

Many researchers and organizations (AAAS, 1989; Abd-El-Khalick, 2005; Abd-El-Khalick & Lederman, 2000; Holbrook & Rannikmae, 2007; NRC, 1996; Schwartz et al., 2004, among others) have argued that science classes should emphasize the nature of science as a fundamental aspect of scientific literacy. They also argue that science should be placed in societal aspects. The nature of science was tested through two protocols, the administration of a survey (VOSE) and an interview using the VNOS-B. The VOSE data indicated that most students were approximately halfway between traditional and

contemporary nature of science views, and these views did not change greatly over the course of the study. There are at least two possible causes for this effect, in addition to the possibility that the SSI curriculum did not have a statistically significant effect on the students' nature of science views. The first possibility is the VOSE survey did not adequately discriminate between naïve and more sophisticated nature of science views. The VOSE was initially developed through a qualitative process, and the use of a five point Likert scale may not have been adequate to discriminate between conceptions regarding the nature of science, which is seen as a developmental process not likely to be changed within a short period of time. A second possible effect is due to the decreased motivation by the students to do the survey a second time for the posttest. One of the teachers during a follow up interview indicated that many students recognized the survey, which surprised the teacher, and some students did not want to complete the survey again.

Understanding of the nature of science has been thought by many (AAAS 1989, 1993; Hodson, 2003; Kolstø, 2001; McComas & Olson, 1998; Symington & Tytler, 2004), to be necessary for preparation as practicing scientists and democratic citizens. In this instance, the incorporation of SSI units without an explicit NOS component did not change NOS understanding greatly over the short term, which would indicate that students' beliefs did not change greatly. The use of an explicit NOS component has been shown to be effective for students with more traditional NOS ideas, while students with transitional views improved with implicit NOS teaching (Khishfe & Lederman, 2006). Although the students (on average) were rated as transitional based on the VOSE scores, perhaps they would have benefited from explicit NOS teaching throughout the course of

the study. However, it should not be inferred that having a sophisticated understanding of the nature of science would lead to decision-making based solely on scientific principles. Bell and Lederman (2003) studied the influence of NOS understanding on adults with varying understanding of NOS principles. They found that NOS experts did not simply utilize scientific principles when making decisions based on scientific issues, rather they used a variety of influences including moral and ethical values, social and political influences, and pragmatism to reach decisions. The NOS expert group did tend to view the evidence to support claims, rather than offer absolute proof, which should be a main goal of NOS instruction rather than expecting students to view issues in purely scientific terms.

2. To what extent do students enrolled in high school science classes utilizing socioscientific issues show greater development in reflective judgment compared to students in traditional high school science classes?

The SSI curriculum is well suited for the ideals of the scientific process, which depends on “empirical standards, logical arguments, and skepticism, as scientists strive for the best possible explanations about the natural world (NRC, 1996, p. 201).” The process of investigating an issue then discussing it with peers holding opposing viewpoints develops the skills necessary to engage in the scientific process. Reflective judgment is well suited to assess these skills, as the model was developed to examine how individuals defend their judgments and how their views of knowledge evolve (King, 2008). Reflective judgment was tested in this study using two different instruments: the RCI computer based test as well as the PRJI protocol. The comparison group decreased from the pretest to posttest, while the treatment group increased slightly over the course of six months.

Although the results were not statistically significant, this may have been an artifact of the study (small number of classes) as the treatment group showed greater increases than would be expected in the literature. As the students taking the RCI included two honors classes and four regular classes with a fair amount of mainstream ESE and low level readers, it could be reasonably inferred that this sample is representative of the high school population. The use of the ISSA methodology, however, decreases the ability to generalize to the general population. The relative success of the reflective judgment portion of the study may be directly related to the similarities that exist between reflective judgment and socioscientific issues: the use of ill-structured problems, the use of evidence to evaluate a position, and the importance of constructed knowledge (Zeidler et al., 2009). Reflective judgment is also a cognitive-developmental construct, which examines people's reasoning about different topics. Although there are many ways to develop critical thinking, King and Kitchener (2002, p. 55) provide some suggestions for including reflective judgment into the curriculum:

1. Show respect for students' assumptions, regardless of the developmental stage(s) they exhibit. Their assumptions are genuine, sincere reflections of their ways of making meaning, and are steps in a developmental progression. If students perceive disrespect or lack of emotional support, they may be less willing to engage in challenging discussions or to take the intellectual and personal risks needed for development.
2. Discuss controversial, ill-structured issues with students throughout their educational activities, and make available resources that show the factual basis and lines of reasoning for several perspectives.

3. Create many opportunities for students to analyze others' points of view for their evidentiary adequacy and to develop and defend their own points of view about controversial issues.
4. Teach students strategies for systematically gathering data, assessing the relevance of the data, evaluating data sources, and making interpretive judgments based on the available data.
5. Give students frequent feedback, and provide both cognitive and emotional support for their efforts.
6. Help students explicitly address issues of uncertainty in judgment-making and to examine their assumptions about knowledge and how it is gained.

Items two through four on this list are parallel to the goals of a well-crafted SSI curriculum, while items one, five, and six on this list discuss the teachers' role in developing reflective judgment in their students. The units used in this study incorporated many of these curricular designs, including the frequent use of ill-structured problems, the opportunity to analyze other's points of view and defend their own views using evidence, and the development of a method to examine their assumptions about knowledge and how it is gained. The students also were given a method to gather information using electronic sources. The teachers in the study provided the support and the classroom climate for the students to take risks during the discussions and encouraged utilizing evidence for their claims. The curriculum and the classroom environment enabled all three treatment classes to increase their average reflective score by 0.34 points, while the comparison classes decreased by an average of 0.12, even though one of the comparison classes posted a 0.23 increase in reflective judgment.

3. To what extent do students who are enrolled in high school science classes utilizing socioscientific issues show greater argumentation skills compared to students in traditional high school science classes?

Informal logic involves the reasoning about premises that are not known. Debate regarding ill-structured problems falls into this category. Although many would argue over the role of argumentation in science as a central practice (Newton et al., 1999), as a mechanism for learning science (Sandoval & Millwood, 2008), and as a mechanism for citizenship education (Zeidler & Sadler, 2008a), there is a good deal of consensus that argumentation should be included in the science education curriculum. Much of the work in argumentation studies involves the structure of the argument rather than the content. This study investigated both the structure of the argument as well as the science content used to support the argument. Many researchers (Erduran et al., 2004; Jimenez-Alexandre et al., 2000; Kelly et al., 2008; Resnick et al., 1993) have found that students often use claims without supporting evidence, and the results from this study echo those results. Those results may be mixed, however, depending on the topic. The essay involving the abortion of a fetus with cystic fibrosis was not conducive to the use of scientific evidence, although it was present in the prompt. Abortion in the United States is a contentious issue, and many people's views on the topic are guided by emotion rather than reason. The water fluoridation prompt did not strike the same emotional cord as the abortion topic, and many students used data from the essay prompt to support their claims. Four of the classes improved from pretest (CF issue) to posttest two (fluoridation), after all six classes decreased on the posttest one (CF issue) essay. Although decreased motivation was one potential factor, the subject matter may have

played a role in students producing low essay scores. Although the scores did not increase greatly, it should be noted that the students had not engaged in any formal debate or argumentation prior to the study. One could argue that while argumentation is a critical part of the SSI protocol, students must be exposed to explicit instruction in the structure of arguments, the use of evidence in constructing arguments, and the role of fallacious reasoning in argumentation.

In addition to the essay data, argumentation was also tested through the use of an argumentation interview protocol. The interviews tested students' ability to develop a position and rationale, take multiple perspectives, utilize a rebuttal, and use subject matter knowledge. Most of the students were able to develop their own position with a short rationale, and many of the students were also able to take the opposite perspective when asked. Very few students, however, were able to provide a rebuttal without simply restating their own position and few students used science subject matter in their answers. This lack of argumentation awareness necessitates the need for explicit argumentation in secondary school. Zeidler and Sadler (2008b) argue that "educational programs and research focused on promoting argumentation and character development should attend to how well students are able to articulate coherent and internally consistent arguments, recognize potential threats to positions and counter positions and form rebuttals." (p. 212) Argumentation, although used throughout the course, was not explicitly taught to the students. Future studies should focus on the combination of explicit argumentation strategies combined with the SSI curriculum to provide a context for those strategies.

Implications for Practice

The use of a socioscientific issues based curriculum was utilized as the primary method of instruction over the course of five units in six heterogeneous Biology I classes in a suburban high school. This treatment was significant as there have been many factors which block the use of argumentation in the classroom. Newton et al. (1999) interviewed fourteen experienced science teachers in England regarding the lack of argumentation in secondary science curriculum. The teachers offered a number of internal and external influences on their teaching priorities. Some of the internal issues were the classroom management skills needed to incorporate argumentation, the lack of quality materials for teachers, teachers' skills and views of science, and the lack of teacher training in the area. Some of the external influences involved the time constraints necessitated by covering a national curriculum for external tests, the students' and parents' views of the need to fill "course books" and students' discomfort with participating in science discussions. As the SSI curriculum is heavily dependent upon the use of argumentation, many of the same influences could be applied to the use of an SSI curriculum. Two different teachers without substantial SSI background were able to facilitate instruction and encourage the students to participate in the class. It was found (Callahan, 2009) that both the teachers and students remarked during interviews that the SSI curriculum enabled students to express opinions, relate science to the real world more effectively than traditional science courses have, and increased student interest in the subject matter. However, the teachers and students both remarked that students sometimes had trouble with the cognitive and behavioral demands placed on the student. Teachers and students reported that literacy skills, particularly in reading and public

speaking would need to be emphasized within the context of the SSI curriculum, as well as the research skills needed to provide the rich background knowledge needed to effectively participate in SSI debates. Some students reported that their classmates were not used to a more interactive classroom, and group dynamics and some student behaviors occasionally disrupted instruction. However, these are some of the same issues faced with any type of cooperative learning, and should not be considered a barrier to utilizing SSI in the classroom. As a result, further classroom practice involving SSI should consider the explicit discussion of basic literacy, including argumentation structure, research skills -- including the use of technology in information technology, and group dynamics in order to provide a classroom environment in which a SSI curriculum is more likely to be successful, particularly with students who have not been exposed to or have had difficulty with such concepts in the past.

One method that could be used to bring some permanence to the SSI curriculum would be the development of an ancillary textbook that incorporates these themes in a more simple to use package for the practicing teacher. Teachers are often faced with large course loads and minimal planning time, and incorporating the various aspects studied here would require the revamping of an entire course. A more prescriptive curriculum that incorporates these themes might alleviate some of the pressure on the teacher to develop an SSI-based (with all it entails) curriculum on his or her own. The packaging of the curriculum with a textbook would provide for stronger links to the content as well as involve the possibility for inquiry based laboratory exercises, which could be used to further support the issues based curriculum.

Two of the constructs measured, nature of science and argumentation, did not show large achievement gains. This finding shows that the length of time of the SSI treatment may not be the sole determinant regarding learning gains. Both of these constructs require a specific mind set (nature of science) or a specific skill set (argumentation), neither of which was explicitly emphasized during the course of this study. Previous research (Khishfe & Lederman, 2006) emphasized the use of integrated and non-integrated NOS instruction as necessary for NOS development. Students with naïve views improved their views using non-integrated NOS instruction, while students with transitional views benefited from integrated NOS instruction. SSI, by its very nature, would be considered integrated NOS instruction, which may not have been helpful to those students with naïve views at the beginning of the study.

Argumentation, particularly the formation of a coherent argument, was also not explicitly taught during this study to prevent the confounding factor of whether argumentation instruction or the SSI curriculum would have been responsible for any learning gains. Although the students were exposed to multiple perspectives on an issue, the students were not significantly better at framing their own arguments than at the beginning of the treatment. This finding provides evidence that argumentation structure and the use of scientific evidence must be explicitly taught to students, and the SSI curriculum provides a potential context from which to teach argumentation.

Reflective judgment, on the other hand, although was not statistically significant, did provide a measure that the SSI treatment was beneficial without any additional instruction. Perhaps the same criteria that make reflective judgment an effective method for evaluating SSI (ill-structured problems, use of evidence to support reasoning, overlap

of topics) also lead to the potential development of reflective judgment through the use of SSI over an extended period of time. In this study we found an average increase within our treatment classes over the course of six months equal to the reported gains in high school students over the course of four years. Although this was a small sample and these findings require further investigation, there is the potential for the development of reflective judgment in high school students through the use of an SSI curriculum.

Limitations of the Study

This goal of this study was to investigate the degree to which the use of a SSI based curriculum affected students' development of three aspects related to scientific literacy: nature of science, reflective judgment, and argumentation. Both qualitative and quantitative data were collected and analyzed with the purpose of investigating this goal. Although there have been some significant findings, the study was not without limitations. One limitation was the relatively small sample size which hindered the ability to find statistical significance for typically slowly developing cognitive constructs. Even with the use of ISSA to more accurately describe the sample set under consideration, statistical significance was not found.

A second limitation involved the time of the day each of the classes met. Although the classes were randomly assigned to either the treatment or the comparison group, the treatment classes tended to cluster at the beginning of the day (starting at 7:00 a.m.) and at the end of the school day (11:20 to 1:05). With these considerations, it is possible that students were not able to perform at their best earlier in the morning and at the end of the day following only a fifteen-minute break between periods four and five.

The lunch period at the school follows the school day, so it is possible that many students had not eaten since breakfast by the time fifth and sixth period science classes were held.

A third limitation of the study was also related to the student population. Due to practical considerations, the sample was taken from two teachers at one high school. One of the teachers had four classes of Biology I, while the second teacher had two classes of Biology I Honors. As these courses were predetermined, there was no method for ensuring they were equivalent prior to the study. Periods one and three contained students that were on average one year older than their counterparts in the other four classes. This might be a factor in the period 3 (comparison group) scoring as the second highest group on all three measures on the posttest, but this hypothesis was not supported by the other older group (period 1, treatment group) which scored the lowest on the posttest for the nature of science and the argumentation essay.

The inclusion of the honors classes did not seem to play a role, as both period two (comparison) and period six (treatment) had similar scores to the other classes. This may have been an artifact of the method of selection for honors classes, namely student and parent choice. If a student signed up for the Honors section of any course, they were given the course, regardless of recommendations or reading and math ability. The result was that student performance in the “honors” classes was often indistinguishable from that of the “regular” sections of the same course.

Period five scored the highest of all classes on all three instruments. This may have been due to the smaller overall class size and the generation of discussion. These factors were not directly measured, however, there is evidence that neither age nor enrollment in an honors section was not a factor for this difference.

Recommendations for Further Research

Increasing the sample size may provide the ability for the researcher to find statistical differences where this study was unable to find them. Using a larger sample size in terms of additional classrooms would increase the ability to infer characteristics from the sample to the larger population of high school students as a whole. The use of additional students would increase the possibility of finding significant differences if they exist.

A second suggestion would be to include explicit non-integrated instruction in nature of science and argumentation structure. Although argumentation typically falls under the language arts curriculum in secondary schools, the use of argumentation accounts for the development of scientific knowledge (Newton, et al., 1999) and is instrumental in the formation of scientific literacy.

A third suggestion would involve a longitudinal study involving SSI as the main science curriculum over a course of three years, the typical number of courses required of high school students. Would the benefits of an SSI curriculum be greater over an extended period of time? The intuitive answer would be affirmative, as the constructs investigated here typically progress slowly, hence the expanded period of time and exposure to these concepts may have the benefit of encouraging a scientific mindset among high school students.

Conclusions

Historically, many (Khishfe & Lederman, 2006, Walker & Zeidler, 2007; Zohar & Nemet, 2002, among others) studies have focused on single unit of instruction treatments. This study is one of the few (Zeidler et al., 2009), however, to utilize a

treatment covering multiple units. Six classes of Biology I were utilized in the study, three receiving five units of SSI instruction, and three classes taught in the normal manner by their teachers. The response to the SSI treatment was generally positive from both the teachers and the students, who reported that the opportunity to express opinions and the connections to the real world, contributed to an increased interest in the subject matter (Callahan, 2009).

This study sought to use a mixed methodology in order to determine both whole group and individual changes in three aspects of scientific literacy: nature of science, reflective judgment, and argumentation. Written surveys and essays were used to gather quantitative data to describe each of the classes, while interviews provided qualitative data to describe individual students' conceptions of scientific literacy.

The findings presented in this study follow the work of many who have contributed greatly to the fields of scientific literacy and socioscientific issues. Although this study found few statistically significant findings, there were indications that in some cases scientific literacy may be enhanced through the incorporation of socioscientific issues in the high school science curriculum, and other cases the SSI curriculum should be combined with explicit instruction in factors deemed important to the teacher or researcher.

Nature of science and argumentation, in particular, were skills that were not explicitly taught during the course of the study, and did not increase greatly. Reflective judgment, on the other hand, did show some increases without explicit instruction; however, it has been argued (Zeidler et al., 2009) that there is a strong correlation between reflective judgment and SSI instruction. The increases in scores may provide

some evidence for that assertion. Qualitative analysis provided examples of students who did progress significantly regarding each of the constructs under investigation, but many of the interviews confirmed much of the evidence from the written instruments, that some students improved throughout the course of the study, but widespread increases were not seen.

The use of the SSI curriculum for a six month period with high school Biology I students provided evidence that SSI could be used to provide a context for science instruction, despite arguments expressed by some science teachers (Newton et al. 1999) regarding the lack of argumentation in secondary science curriculum. The teachers offered a number of influences on their teaching priorities including classroom management skills, lack of quality materials, teachers' skills and views of science, and the lack of teacher training. However, the evidence does not support the argument of a lack of quality materials. With the expansion of the Internet, there are a number of high quality science units available to the teacher. The other issues presented should be addressed in teacher education programs. Classroom management, content knowledge, and pedagogical knowledge are all issues that need to be addressed by educators of pre-service teachers and were outside the scope of the current study.

There is a gap between science education research and practice. While the research indicates that learning science involves inquiry and analysis, many teachers present science in a traditional, fact-based manner. Consequently, what students learn in the classroom is disconnected from their daily lives (Duit & Treagust, 1998, National Research Council, 1996, 2000). The use of a SSI based curriculum directly addresses the

issue of personal relevance to science and may provide the context for the development of many tenets of scientific literacy.

REFERENCES

- Abd-El-Khalick, F. (2005). Developing deeper understandings of nature of science: the impact of a philosophy of science course on preservice science teachers' views and instructional planning. *International Journal of Science Education*, 27, 15-42.
- Abd-El-Khalick, F. (2003). Socioscientific issues in pre-college science classrooms: The primacy of learners' epistemological orientations and views of nature of science. In D.L. Zeidler (Ed.) *The role of moral reasoning on socioscientific issues and discourse in science education*. Dordrecht: Kluwer Academic Press.
- Abd-El-Khalick, F. (1998). The influence of history of science courses on students' conceptions of the nature of science. Unpublished doctoral dissertation, Oregon State University, Oregon.
- Abd-El-Khalick, F., Bell, R.L., & Lederman, N.G. (1998). The nature of science and instructional practice: Making the unnatural natural. *Science Education*, 82, 417-436.
- Abd-El-Khalick, F., & Lederman, N.G. (2000). Improving science teachers' conceptions of nature of science: a critical review of the literature. *International Journal of Science Education*, 22, 665-701.
- Aikenhead, G.S., Ryan, A.G. & Desautels, J. (1989, April). Monitoring student views on science-technology-society issues: the development of multiple choice items. Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.
- Aikenhead, G., Ryan, A., & Fleming, R. (1989). *Views on science-technology-society (from CDN.mc.5)*. Saskatoon, Canada: Department of Curriculum Studies, University of Saskatchewan.
- American Academy for the Advancement of Science (1993). *Benchmarks for Science Literacy: a Project 2061 Report*. New York: Oxford University Press.
- American Academy for the Advancement of Science (1989). *Project 2061: Science for All Americans*. Washington DC: Author
- Bell, R.L. (1999). Understandings of the nature of science and decision making on science and technology based issues. Unpublished doctoral dissertation, Oregon State University, Oregon.

- Bell, R.L. & Lederman, N.G. (2003). Understandings of the nature of science and decision making on science and technology based issues. *Science Education*, 87, 352-377.
- Bell, P. & Linn, M.C. (2000). Scientific arguments as learning arguments: designing for learning from the web with KIE. *International Journal of Science Education*, 22(8), 797-817.
- Berkowitz, M.W., Oser, F. & Althof, W. (1987). The development of sociomoral discourse. In W.M. Kurtines & Gewirtz, J.L. (Eds.), *Moral development through social interaction*. New York: John Wiley.
- Berkowitz, M.W. & Simmons, P. (2003). Integrating science education and character education: The role of peer discussion. In D.L. Zeidler (ed.) *The role of moral reasoning on socioscientific issues and discourse in science education*. Dordrecht: Kluwer Academic Press.
- Broughton, J.M. (1978). Development of concepts of self, mind, reality, and knowledge. In W. Damon (Ed.), *Social cognition: New directions for child development*, no. 1. San Francisco: Jossey-Bass.
- Burek, K., Callahan, B., & Zeidler, D.L. (2004, October) Argumentation for Scientific Literacy: Seal Hunting in Canada. Paper presented at the Annual Meeting for the Southern Association for the Education of Teachers of Science, Gainesville, FL.
- Bybee, R.N. (1997). Towards an understanding of scientific literacy. In W. Grabe and C. Bolte (Eds.) *Scientific Literacy – An International Symposium*, IPN, Kiel, Germany.
- Callahan, B.E. (2009, April). Teacher and student perceptions regarding the incorporation of socioscientific issues in high school biology. Paper presented at the Annual Meeting of the National Association for the Research in Science Teaching national conference, Garden Grove, CA.
- Callahan, B.E., Robinson, C., & Fowler, S. (2006, April). Literacy with a Purpose: Linking Science, Literacy, and Technology. Paper presented at the National Science Teachers Association national conference, Anaheim, CA
- Callahan, B.E., Zeidler, D.L., Cone, N. & Burek, K. (2005, October). The effects of learning socioscientific issues on reflective judgment in high school science students. Paper presented at the Annual Meeting of the Southeastern Association for Science Teacher Education, Athens, GA.
- Chen, S. (2006a). Development of an instrument to assess views on nature of science and attitudes toward teaching science. *Science Education*, 90, 803-819.

- Chen, S. (2006b). Views on science and education (VOSE) questionnaire. *Asia-Pacific Forum on Science Learning and Teaching*, 7:2, article 11.
- Cowlshaw, R., Hunter, C., Coy, J., & Tessmer, M. (2006). The art of a deal: a Kyoto Protocol simulation. Published by the National Center for Case Study Teaching in Science http://www.sciencecases.org/kyoto_protocol/kyoto_protocol.asp
- Davies, I. (2004). Science and citizenship education. *International Journal of Science Education*, 26(14), 1751-1763.
- Dotger, S. & Jones, M.G. (2007, April). Reflective Judgment and Nature of Science: Commonalities Explored. Paper presented at the Annual Meeting for the National Association for Research in Science Teaching, New Orleans, LA.
- Driver, R., Newton, P. & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84(3), 287-312.
- Duit, R. & Treagust, D.F. (1998). Learning in science – from behaviourism towards social constructivism and beyond. In Fraser, B.J. & Tobin, K.G. (eds.) *International Handbook of Science Education* (p. 3-25). Great Britain: Kluwer.
- Duschl, R.A. (2008). Quality argumentation and epistemic criteria. In Erduran, S. & Jimenez-Aleixandre, M.P. (eds.), *Argumentation in Science Education* (p. 159-175). Springer
- Duschl, R.A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, 38, 39-72.
- Erduran, S., Simon, S., & Osborne, J. (2004). TAPPING into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, 88, 915-933.
- Felton, M.K. (2004). The development of discourse strategies in adolescent argumentation. *Cognitive Development*, 19, 35-52.
- Feretti, R.P., Andrews-Weckerly, S. & Lewis, W.E. (2007). Improving the argumentative writing of students with learning disabilities: descriptive and normative considerations. *Reading and Writing Quarterly*, 23(3), 267-285.
- Fischer, K.W. (1980). A theory of cognitive development: the control and construction of hierarchy of skills. *Psychological Review*. 87(6), 477-531.
- Fowler, S.R., Zeidler, D.L., Sadler, T.D. (2009). Moral Sensitivity in the Context of Socioscientific Issues in High School Science Students. *International Journal of Science Teacher Education*, 31(2), 279-296.

- Giere, R.N. (1988). *Exploring science: A cognitive approach*. Chicago: The University of Chicago Press.
- “Grades 4, 8 post big gains on NAEP writing test. (National Assessment of Educational Progress).” *Education USA* 45.8 (August 2003): 10(1).
- Hodson, D. (2003). Time for action: science education for an alternative future. *International Journal of Science Education*, 25, 645-670.
- Iding, M. & Klemm, E.B. (2005). Pre-service teachers critically evaluate scientific information on the world wide web: what makes information believable? *Computers in the Schools*, 22(1-2), 7-18.
- Holbrook, J. & Rannikmae, M. (2007). The nature of science education for enhancing scientific literacy. *International Journal of Science Education*, 29, 1347-1362.
- Jimenez-Aleixandre, M.P., Rodriguez, A.B., & Duschl, R.A. (2000). “Doing the lesson” or “doing science”: argumentation in high school genetics. *Science Education*, 84, 757-792.
- Kelly, G.J., Regev, J., & Prothero, W. (2008). Analysis of lines of reasoning in written argumentation. In Erduran, S. & Jimenez-Aleixandre, M.P. (eds.), *Argumentation in Science Education* (p. 137-157). Springer
- King, P.M. (accessed 3/24/08). Reflective judgment: Welcome to the website. <http://www.umich.edu/~refjudg/index.html>
- King, P.M. & Kitchener, K.S. (1994). *Developing Reflective Judgment: Understanding and Promoting Intellectual Growth and Critical Thinking in Adolescents and Adults*. San Francisco: Jossey-Bass.
- King, P.M. & Kitchener, K.S. (2002). The reflective judgment model: Twenty years of research on epistemic cognition. In B.K. Hofer & P.R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*. (pp. 37-61). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- King, P.M. & Kitchener, K.S. (2004). Reflective judgment: theory and research on the development of epistemic assumptions through adulthood. *Educational Psychologist*, 39, 5-18.
- Kitchener, K.S. (1983). Educational goals and reflective thinking. *Educational Forum*, 48(1), 75-95.
- Kitchener, K.S., King, P.M., Wood, P.K., & Davison, M (1989). Sequentiality and consistency in the development of reflective judgment: a six-year longitudinal study. *Journal of Applied Psychology*, 10, 13-95.

- Khishfe, R. & Lederman, N. (2006). Teaching nature of science within a controversial topic: integrated versus nonintegrated. *Journal of Research in Science Teaching*, 43, 395-418.
- Kolstø, S.D. (2001). Scientific literacy for citizenship: Tools for dealing with the science dimension of controversial socioscientific issues. *Science Education*, 85, 291-310.
- Kolstø, S.D., Bungum, B., Arnesen, E., Isnes, A., Kristensen, T., Mathiassen, K., Mostad, I., Quale, A., Sissal Vedvik Tonning, A., & Ulvik, M. (2006). Science student's critical examination of scientific information related to socioscientific issues, *Science Education*, 90, 632-655.
- Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education*, 77(3), 319-337.
- Kuhn, D. (1991). *The Skills of Argument*. New York: Cambridge University Press.
- Kuhn, L. & Reiser, B. (2005). Students constructing and defending evidence-based scientific explanations. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Dallas, TX.
- Kuhn, T.S. (1962/1996) *The Structure of Scientific Revolutions*. 3rd edition. Chicago: University of Chicago Press.
- Lederman, N.G. (2007). Nature of science: past, present, and future. In S.K. Abell and N.G. Lederman (Eds.). *Handbook of Research on Science Education*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Lederman, N.G., Abd-El-Khalick, F., Bell, R. & Schwartz, R.S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conception of the nature of science. *Journal of Research in Science Teaching*, 39, 497-521.
- Lederman, N.G. & O'Malley, M. (1990). Students' perceptions of tentativeness in science: development, use, and sources of change. *Science Education*, 74, 225-239.
- Lederman, N.G., Wade, P.D., & Bell, R.L. (1998). Assessing understanding of the nature of science: A historical perspective. In McComas, W. (Ed.), *The nature of science in science education: Rationales and strategies* (pp. 331-350). The Netherlands: Kluwer Academic.
- Lincoln, Y.S. & Guba, E.G. (1985). *Naturalistic Inquiry*. Newbury Park, CA: Sage Publications.

- Marttunen, M., Laurinen, L., Litosselitti, L., & Lund, K. (2005). Argumentation skills as prerequisites for collaborative learning among Finnish, French, and English secondary school students. *Educational Research and Evaluation*, 11(4), 365-384.
- McComas, W.F., Clough, M.P., & Almazroa, H. (1998). The role and character of the nature of science in the science education. In W.F. McComas (Ed.) *The nature of science in science education: Rationales and strategies*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- McComas, W.F. & Olson, J.K. (1998). The nature of science in international science education standards documents. In W.F. McComas (Ed.) *The nature of science in science education: Rationales and strategies*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- National Center for Case Study Teaching in Science (Producer). (2002). The use of case studies and group discussions in science education. [Instructional video]. Available from the National Center for Case Study Teaching in Science, c/o Carolyn Wright, Dept. of Biological Sciences, University of Buffalo, 109 Cooke Hall, Buffalo, NY 14260.
- National Center for Education Statistics (2004). Highlights from the trends in international mathematics and science study (TIMSS) 2003. Washington, DC: U.S. Department of Education.
- National Research Council (2000). *Inquiry and the National Science Education Standards*. Washington D.C.: National Academy Press
- National Research Council (1996). *National Science Education Standards*. Washington DC: National Academy Press.
- Newton, P., Driver, R. & Osborne, J. (1999). The place of argumentation in the pedagogy of school science. *International Journal of Science Education*, 21, 553-576.
- Norris, S.P. & Phillips, L.M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, 87, 224-240
- Organisation for Economic Co-operation and Development (2006). *PISA 2006: Science competencies for tomorrow's world*. Vol. 1.
- Organisation for Economic Co-operation and Development (2003). *The PISA 2003 assessment framework: Mathematics, reading, science and problem solving knowledge and skills*. Eric Reproduction Service No. ED 480801.

- Organisation for Economic Co-operation and Development (1998). Instrument design: a framework for assessing scientific literacy. Report of project managers meeting. Arnhem: Programme for International Student Assessment.
- Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What “ideas about science” should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching*, 40, 692-720.
- Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in science classrooms. *Journal of research in science teaching*, 41(10), 994-1020.
- Parendes, L.A., & Burris, S.H. (2005). Pesticides: Can we do without them? Published by the National Center for Case Study Teaching in Science
<http://www.sciencecases.org/pesticides/pesticides.asp>
- Pella, M.D., O’Hearn, G.T., & Gale, C.W. (1966). Referents to scientific literacy. *Journal of research in science teaching*, 4, 199-208.
- Perry, W.G. (1970). *Forms of Intellectual and Ethical Development in the College Years: A Scheme*. New York: Holt, Rinehart, and Winston.
- Piaget, J. & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books.
- Popper, K. (1975/1998). The rationality of science revolutions. In J.A. Kourany (Ed.), *Scientific knowledge* (pp. 286-300), Wadsworth, CA: Belmont. (Reprinted from *Problems of scientific revolution: progress and obstacles to progress in the sciences*, pp. 72-101, by R. Harre, Ed., 1975, Oxford: Clarendon Press.)
- Resnick, L.B., Salmon, M., Zeitz, C.M., Wathen, S.H., & Holowchak, M. (1993). Reasoning in conversation. *Cognition and Instruction*, 11(3&4), 347-364.
- Roberts, D.A. (2007). Scientific literacy / science literacy. In S.K. Abell and N.G. Lederman (Eds.) *Handbook of research on science education*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Rubin, L.M. (2003). Saving Superman: A look into stem cell research. Published by the National Center for Case Study Teaching in Science
<http://www.sciencecases.org/superman/superman1.asp>
- Rudolph, J.L. (2003). Portraying epistemology: school science in historical context. *Science Education*, 87, 64-79.
- Ryan, M.E. (2001). The Tokaimura Accident: Nuclear energy and reactor safety. Published by the National Center for Case Study Teaching in Science
<http://ublib.buffalo.edu/libraries/projects/cases/tokaimura/tokaimura.html>

- Sadler, T.D. (2006). Promoting discourse and argumentation in science teacher education. *Journal of Science Teacher Education*, 17(4), 323-346.
- Sadler, T.D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching*, 41(5), 513-536.
- Sadler, T.D. (2002). Socioscientific issues and the affective domain: Scientific literacy's missing link. (Report No. SE-067-119). Paper presented at the Annual Meeting of the Southeastern Association for the Education of Teachers in Science, Kennesaw, GA. (ERIC Document Reproduction Service No. ED472041).
- Sadler, T.D. Chambers, F.W., & Zeidler, D.L. (2004). Student conceptualizations of the nature of science in response to a socioscientific issue. *International Journal of Science Education*, 26, 387-409.
- Sadler, T.D., & Donnelly, L.A. (2006). Socioscientific argumentation: the effects of content knowledge and morality. *International Journal of Science Education*, 28, 1463-1488.
- Sadler, T.D. & Fowler, S.R. (2006). A threshold model of content knowledge transfer for socioscientific argumentation. *Science Education*, 90(6), 986-1004.
- Sandoval, W.A. & Millwood, K.A. (2008). What can argumentation tell us about epistemology? In Erduran, S. & Jimenez-Aleixandre, M.P. (eds.), *Argumentation in Science Education* (p. 71-88). Springer
- Sandoval, W.A. & Millwood, K.A. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition and Instruction*, 23(1), 23-55.
- Schwartz, R.S., Lederman, N.G., & Crawford, B.A. (2004). Developing views of nature of science in an authentic context: an explicit approach to bridging the gap between nature of science and scientific inquiry. *Science Education*, 88, 610-645.
- Scott, P. (1998). Teacher talk and meaning making in science classrooms: A Vygotskian analysis and review. *Studies in Science Education*, 32, 45-80.
- Shaffer, D.W. & Serlin, R.C. (2004). What good are statistics that don't generalize? *Educational Researcher*, 33(9), 14-25.
- Shakirova, D.M. (2007). Technology for the shaping of college students' and upper-grade students' critical thinking. *Russian Education and Society*, 49(9), 42-52.
- Shew, W. & Reese, M.C. (2007). Do you really know what you're eating? A case study on genetically modified foods. Published by the National Center for Case Study Teaching in Science http://www.sciencecases.org/gmo_safety/gmo_safety.asp

- Sipress, J.M. (2004). Why students don't get evidence and what we can do about it. *History Teacher*, 37(3), 351-363.
- Stevens, J. (1999). *Intermediate Statistics: A modern approach*. 2nd edition. Mahwah, NJ: Lawrence Erlbaum Associates.
- Symington, D. & Tytler, R. (2004). Community leader's views of the purposes of science in the compulsory years of teaching. *International Journal of Science Education*, 26(11), 1403-1418.
- Teddlie, C. & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In A. Tashakkori and C. Teddlie (Eds.) *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage Publications.
- Toulmin, S.E. (1958/2003). *The uses of argument*. Updated edition. New York: Cambridge University Press.
- Waghid, Y. (2005). Action as an educational virtue: toward a different understanding of democratic citizenship education. *Educational Theory*, 55(3), 323-342.
- Walker, K.A. & Zeidler, D.L. (2007). Promoting discourse about socioscientific issues through scaffolded inquiry. *International Journal of Science Education*, 29(11), 1387-1410.
- Watson, J.R., Swain, J.R.L., & McRobbie, C. (2004). Students' discussions in practical scientific inquiries. *International journal of science education*, 26(1), 35-45.
- Weinstock, M., Neuman, Y., & Tabak, I. (2004). Missing the point or missing the norms? Epistemological norms as predictors of students' ability to identify fallacious arguments. *Contemporary Educational Psychology*, 29(1), 77-94.
- Yore, L.D., Florence, M.K., Pearson, T.W., & Weaver, A.J. (2006). Written discourse in scientific communities: A conversation with two scientists about their views of science, use of language, role of writing in doing science, and compatibility between their epistemic views and language. *International Journal of Science Education*, 28, 109-141.
- Zeidler, D.L. (2007, May). An inclusive view of scientific literacy: Core issues and future directions. Paper presented at the Linnaeus Tercentenary 2007 Symposium "Promoting Scientific Literacy," Uppsala University, Uppsala, Sweden.
- Zeidler, D.L., Applebaum, S. & Sadler, T.D. (2006, January). Using socioscientific issues as a context for teaching content and concepts. Paper presented at the Annual Meeting of the Association for Science Teacher Education, Portland, OR.

- Zeidler, D.L. & Keefer, M. (2003). The role of moral reasoning and the status of socioscientific issues in science education: Philosophical, psychological, and pedagogical considerations. In D.L. Zeidler (Ed.) *The role of moral reasoning and discourse on socioscientific issues in science education* (pp. 7-38). Netherlands: Kluwer.
- Zeidler, D.L. & Sadler, T.D. (2008a). Social and ethical issues in science education: A prelude to action. *Science & Education*, 17(8,9).
- Zeidler, D.L. & Sadler, T.D. (2008b). The role of moral reasoning in argumentation: Conscience, character and care. In S. Erduran & M. Pilar Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research*. The Netherlands: Springer Press.
- Zeidler, D.L., Sadler, T.D., Applebaum, S., & Callahan, B.E. (2009). Advancing reflective judgment through socioscientific issues. *Journal of Research in Science Teaching*, 46, 74-101.
- Zeidler, D.L., Sadler, T.D., Simmons, M.L. & Howes, E.V. (2005). Beyond STS: A research-based framework for socioscientific issues education. *Science Education*, 89, 357-377.
- Zeidler, D.L., Walker, K.A., Ackett, W.A. & Simmons, M.L. (2002). Tangled up in views: beliefs in the nature of science and responses to socioscientific dilemmas. *Science Education*, 86(3), 343-367.
- Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, 39, 35-62.

APPENDICES

APPENDIX A: VIEWS ON SCIENCE AND EDUCATION SURVEY

Views on Science and Education Questionnaire

Instructions to participants:

Each question of this questionnaire starts with a statement about the nature of science or science education. Most statements adopt a certain radical stance. You may strongly agree with it, strongly disagree with it, or have other thoughts about it. Each statement is followed by several responses. Please read all of the responses, first, then circle your opinion on the right side (SD, D, U, A, SA) of each response according to your knowledge of scientific activities or scientists, or what ought to be taught in science courses. There is no right or wrong answer. Thank you.

SD = Strongly Disagree

D = Disagree

U = Uncertain or No Comment

A = Agree

SA = Strongly Agree

1. When two different theories arise to explain the same phenomenon (e.g., fossils of dinosaurs), will scientists accept the two theories at the same time?

A. Yes, because scientists still cannot objectively tell which one is better, therefore, they will accept both tentatively.	SD	D	U	A	SA
B. Yes, because the two theories may provide explanations from different perspectives, there is no right or wrong.	SD	D	U	A	SA
C. No, because scientists tend to accept the theory are more familiar with.	SD	D	U	A	SA
D. No, because scientists tend to accept the simpler theories and avoid complex theories.	SD	D	U	A	SA
E. No, the academic status of each theory proposer will influence scientists' acceptance of the theory.	SD	D	U	A	SA
F. No, scientists tend to accept new theories which deviate less from the contemporary core scientific theory.	SD	D	U	A	SA
G. No, scientists use intuition to make judgments.	SD	D	U	A	SA
H. No, because there is only one truth, scientists will not accept any theory before distinguishing which is best.	SD	D	U	A	SA

Appendix A: Continued

2. Scientific investigations are influenced by socio-cultural values (e.g., current trends, values).

A. Yes, socio-cultural values influence the direction and topics of scientific investigations.	SD	D	U	A	SA
B. Yes, because scientists participating in scientific investigations are influenced by socio-cultural values.	SD	D	U	A	SA
C. No, scientists with good training will remain value-free when carrying out research.	SD	D	U	A	SA
D. No, because science requires objectivity, which is contrary to the subjective socio-cultural values.	SD	D	U	A	SA

3. When scientists are conducting scientific research, will they use their imagination?

A. Yes, imagination is the main source of innovation.	SD	D	U	A	SA
B. Yes, scientists use their imagination more or less in scientific research.	SD	D	U	A	SA
C. No, imagination is not consistent with the logical principles of science.	SD	D	U	A	SA
D. No, imagination may become a means for a scientist to prove his point at all costs.	SD	D	U	A	SA
E. No, imagination lacks reliability.	SD	D	U	A	SA

4. Even if the scientific investigations are carried out correctly, the theory proposed can still be disproved in the future.

A. Scientific research will face revolutionary change, and the old theory will be replaced.	SD	D	U	A	SA
B. Scientific advances cannot be made in a short time. It is through a cumulative process; therefore, the old theory is preserved.	SD	D	U	A	SA
C. With the accumulation of research data and information, the theory will evolve more accurately and completely, not being disproved.	SD	D	U	A	SA

Appendix A: Continued

5. Is scientific theory (e.g., natural selection, atomic theory) “discovered” or “invented” by scientists from the natural world?

A. Discovered, because the idea was there all the time to be uncovered.	SD	D	U	A	SA
B. Discovered, because it is based on experimental facts .	SD	D	U	A	SA
C. Some scientists discover a theory accidentally, but other scientists may invent a theory from their known facts.	SD	D	U	A	SA
D. Invented, because a theory is an interpretation of experimental facts , and experimental facts are discovered by scientists.	SD	D	U	A	SA
E. Invented, because a theory is created or worked out by scientists.	SD	D	U	A	SA
F. Invented, because a theory can be disproved.	SD	D	U	A	SA

6. Is scientific law (e.g., gravitational law) “discovered” or “invented” by scientists from the natural world?

A. Discovered, because scientific laws are out there in nature, and scientists just have to find them.	SD	D	U	A	SA
B. Discovered, because scientific laws are based on experimental facts .	SD	D	U	A	SA
C. Some scientists discover a law accidentally, but other scientists may invent a law from their known facts.	SD	D	U	A	SA
D. Invented, because scientists invent scientific laws to interpret discovered experimental facts.	SD	D	U	A	SA
E. Invented, since there are no absolutes in nature, therefore, the law is invented by scientists.	SD	D	U	A	SA

7. In comparison to laws, theories have less evidence to support them.

A. Yes, theories are not as definite as laws.	SD	D	U	A	SA
B. Yes, if a theory stands up to many tests it will eventually become a law, therefore, a law has more supporting evidence.	SD	D	U	A	SA
C. Not quite, some theories have more supporting evidence than some laws.	SD	D	U	A	SA
D. No, theories and laws are different types of ideas. They cannot be compared.	SD	D	U	A	SA

Appendix A: Continued

8. Scientists’ observations are influenced by personal beliefs (e.g., personal experiences, presumptions); therefore, they may not make the same observations for the same experiment.

A. Observations will be different, because different beliefs lead to different expectations influencing the observation.	SD	D	U	A	SA
B. Observations will be the same, because the scientists trained in the same field hold similar ideas.	SD	D	U	A	SA
C. Observations will be the same, because through scientific training scientists can abandon personal values to conduct objective observations.	SD	D	U	A	SA
D. Observations will be the same, because observations are exactly what we see and nothing more. Facts are facts. Interpretations may be different from one person to another, but observations should be the same.	SD	D	U	A	SA
E. Observations will be the same. Although subjectivity cannot be completely avoided in observation, scientists use different methods to verify the results and improve objectivity.	SD	D	U	A	SA

Appendix B

VNOS-B Questionnaire

1. After scientists have developed a theory (e.g. atomic theory), does the theory ever change? If you believe that theories do change explain why we bother to teach scientific theories? Defend your answer with examples.
2. What does an atom look like? How certain are scientists about the structure of the atom? What specific evidence do you think scientists used to determine what an atom looks like?
3. Is there a difference between a scientific theory and a scientific law? Give an example to illustrate your answer.
4. How are science and art similar? How are they different?
5. Scientists perform experiments/investigations when trying to solve problems. Other than the planning and the design of these experiments/investigations, do scientists use their creativity and imagination during and after data collection? Please explain your answer and provide examples if appropriate.
6. Is there a difference between scientific knowledge and opinion? Give an example to illustrate your answer.
7. Some astronomers believe that the universe is expanding while others believe it is shrinking; still others believe that the universe is in a static state without any expansion or shrinkage. How are these different conclusions if all of these scientists are looking at the same experiments and data?

The Reasoning about Current Issues Test

(Copyright 2000, Karen Kitchener, Patricia King, & Phillip Wood, all rights reserved)

Page 1 of 2

The Reasoning about Current Issues Test

Demographic and Academic Information

- I. Student ID Number: _____
- II. Birthdate: _____ MM/DD/YY
- III. Are You (check one) Female
 Male
- IV. If you can recall, please provide:
Your ACT composite score: _____
Your ACT composite percentile rank: _____
Your SAT Total score (Verbal + Quantitative) _____
Your SAT percentile rank: _____

V. Racial/Ethnic Classification:

- American Indian/Native American
 Asian
 B l a c k
 Hispanic/Latino/Latina
 White/Caucasian
 International Student
 Other: Specify:

VI. Based on the number of current credit hours toward your degree, would you describe yourself as a:

- Freshman
 Sophomore
 Junior
 Senior
 Beginning Graduate Student (having completed less than three years of graduate coursework)
 Advanced Graduate Student (having completed three or more years of graduate coursework)

Part II: Reasoning About Current Issues

Instructions: Because this questionnaire is aimed at understanding how people like you think about various current issues, it asks not only what you think but why you hold the opinions you do.

The Task: You will be shown five short descriptions of some current issues. These issues are similar because people sometimes disagree about the best answer. For each issue, you will be asked consider four general questions.

Appendix C: Continued

Question 1: In Question 1, you will be asked for your personal opinion about the issue. Please indicate it in the space provided.

Question 2: For some issues you will be asked:
Why experts disagree.

For other issues you will be asked:
Why you believe the way you do.

Take a moment to consider your opinion about the question. Write down your response to the question in a few sentences in the space provided. (Do not, for example, write down "I think experts disagree." or "I think that food additives are safe." Instead indicate in a few sentences why experts disagree or you believe the way you do.

Please give the best answer you have to each question,

Question 3. You will be shown statements taken from interviews with people like yourself.

Please Indicate which statements are most similar to your own views by darkening the appropriate square.

Boxes VS, S, D, and VD are used to indicate whether your response is Very Similar, Similar, Dissimilar, or Very Dissimilar to your own thinking.

For example, if you read sentence A below and decided that it was similar to your views, you would darken the box labeled **S** as follows:

(VS) (■) (D) (VD) (M) A. Researchers who are honest will not disagree about whether a particular artificial sweetener is harmful.

It may be that your views on a topic do not exactly match the ones presented here. Please indicate a few statements for each issue which are at least somewhat similar.

A Check on Reading: Because we have found that some people do not read the statements carefully, we have included some statements that should not make sense to you. When you encounter such statements, mark them as "Meaningless" by darkening the (M).

Question 4. You will be asked to indicate your first, second, and third choices for which statements are like how you think.

Appendix C: Continued

(VS) (S) (D) (VD) (M) F. Researchers arrive at different conclusions because the evidence itself is complex and they examine it from several perspectives. They arrive at a decision by synthesizing their knowledge, experiences, and expert opinions.

(VS) (S) (D) (VD) (M) G. Researchers might say that one view about the safety of a sweetener was better,
but they would also say that this viewpoint is relative to a particular way of understanding- this issue.

(VS) (S) (D) (VD) (M) H. Researchers disagree because the premeditated hard evidence is synthesized
into available belief systems about different comprehensive factual analyses.

(VS) (S) (D) (VD) (M) I. Researchers disagree because they are really studying different facets of the
issue and the best ways to address one facet of the issue are different than the best ways to address other facets.

(VS) (S) (D) (VD) (M) J. Researchers disagree because their evaluation of the evidence leads them to
defend different conclusions. Some researcher's conclusions are more reasonable, however, and reflect a more comprehensive synthesis of the available information.

4. Please rank the statements above (A, B, C., etc.) that are most similar to your thinking. Please check only one statement per line. If no statement beyond one or two is at all like your thinking, check the box labeled "None of These" on the appropriate line(s).

Statement (A) (B) (C) (D) (E) (F) (G) (H) (I) (J) is most like how I think.

Statement (A) (B) (C) (D) (E) (F) (G) (H) (I) (J) [None of these] is second most like how I think.

Statement (A) (B) (C) (D) (E) (F) (G) (H) (I) (J) [None of these] is third most like how I think

Appendix D

Prototypic Reflective Judgment Interview protocol – issues and probe questions (adapted

from King & Kitchener, 1994, p. 100-103; 260; Zeidler, et al., 2009)

Interview Protocol Methodology

“During this session, we will be talking about several issues that are of general concern and about which most people are at least vaguely familiar. I am not concerned with *how much information* you have about any issue, but *how you think* about them. In order to standardize what we talk about, I will be asking the same series of questions for each of the three issues; I am not repeating the questions because I am looking for a particular answer. For each issue, I will read a statement aloud while you follow along on a card. After I finish reading the statement, I’ll give you a minute or so to think about the issue and then we will talk about it. Are there any questions before we begin?”

(Inform participant of the fact that the interview will be tape-recorded. Give the participant a copy of the story to read to him or herself as you read it aloud.)

Reflective Judgment Issues

1) Chemical Additives Issue

There have been frequent reports about the relationship between chemical that are added to foods and the safety of these foods. Some studies indicate that such chemical can cause cancer, making these foods unsafe to eat. Other studies, however, show that chemical additives are not harmful, and actually make the foods containing them more safe to eat.

2) Religion & Science Issue

Many religions of the world have creation stories. These stories suggest that a divine being created the earth and its people. Scientists claim, however, that people evolved from lower animal forms (some of which were similar to apes) into the human forms known today.

3) Alcoholism Issue

Some researchers contend that alcoholism is due, at least in part, to genetic factors. They often refer to results from a number of family studies to support this contention. Other researchers, however, do not think that alcoholism is in any way inherited. They claim that alcoholism is psychologically determined. They also claim that the reason that several members of the same family often suffer from alcoholism is due to the fact that they share common family experiences, socioeconomic status, or employment.

Field Caveat 1. If the response to a given probe question (below) is incomplete, ambiguous, or contradictory to earlier statements, ask for clarification or elaboration by asking such questions as:

- “What do you mean by ‘the most reasonable’ explanation?”
- “What’s the difference between ‘knowing for sure’ and ‘being fairly certain’?”
- “How did that experience change your thinking about [x]?”

Field Caveat 2. If the participant replies that experts disagree because they had access to different information, or held different personal beliefs, then ask:

Appendix D: Continued

- “What if they both had access to the *same* information?”
- “What if they both believed in God [or were both atheists] and still held different views about whether or not evolution occurred?”

Reflective Judgment Interview Standard Probe Questions

Probe Question	Purpose
1. What do you think about these statements? (Note: If no particular point of view is endorsed, ask: 1a) Could you ever say which was the better position? How? Why not? How would you go about making a decision about this issue? Will we ever know for sure which is the better position? How/Why not?	To allow participant to share an initial reaction to the problem presented. Most state which point of view is closer to their own.
2. How did you come to hold that point of view?	To find out how the respondent arrived at the point of view, and whether and how it has evolved from other positions on the issue.
3. On what do you base that point of view?	To find out about the basis of the respondent’s point of view, such as a personal evaluation of the data, consistency with an expert’s point of view, or a specific experience This provides information about the respondent’s concept of justification.
4. Can you ever know for sure that your position on this issue is correct? How or why not?	To find out about assumptions concerning the certainty of knowledge (e.g. whether issues like this can be known absolutely and what the respondent would do in order to increase the certainty, or why that would not be possible.
5. When two people differ about matters such as this, is it the case that one opinion is right and one is wrong? If yes, what do you mean by “right”? If no, can you say that one opinion is in some way better than the other? What do you mean by better”?	Assesses the adequacy of alternative interpretations; to see if dichotomous either/or view of the issue (characteristic of the early stages) is held; to allow the participant to give criteria by which she or he evaluates the adequacy of arguments (information that helps differentiate high-from middle-level stage responses).
6. How is it possible that people have such different points of view about this subject?	To elicit comments about the respondent’s understanding of differences in perspectives and opinions (what they are based on and why there is such diversity of opinion about the issue).
7. How is it possible that experts in the field disagree about this subject?	To elicit respondent’s understanding of how he or she uses the point of view of an expert or authority in making decisions about controversial issues (such as whether experts’ views are weighted more heavily than others’ views, and why or why not).

Appendix E: Description of SSI units used in study

Unit one: Tokaimura Accident.

Part I of this unit relays some background material behind Japan's dependence on nuclear power, as well as the science content behind the theory of nuclear reactions.

Questions at the end of this section ask the students to understand the process of nuclear fission, as well as investigate thoughts about alternative forms of energy.

Part II of this unit provides the chronology of the accident at Tokaimura, along with some questions designed to gauge understanding of the situation, as well as engage students in thinking about possible alternatives that could have prevented the accident from occurring.

Part III details the effects of radiation exposure, as well as the terminology used to discuss irradiation of humans. The questions at the end of this section are designed to gauge understanding of the nature of radiation, and emphasize an understanding that radiation is found naturally.

Part IV deals with the aftermath of the accident. The scenario of what actually happened is contrasted with discussion questions about whether company and government officials acted appropriately following the accident, as well as places the students in different roles as people influenced by the accident.

Part V provides an update of the situation at Tokaimura, as well as asks the students to answer questions regarding the appropriateness of the company and government as it relates to the class discussion from Part IV of the unit. The students also discuss the responsibility of multiple people involved in the accident.

Appendix E: Continued

Part VI provides the students with the outcomes of an independent investigation into the accident, and places the students in the role of the investigators charged with making recommendations for the future to prevent this type of accident of occurring again.

Part VII provides some historical perspective of nuclear power, including information about other famous accidents, such as the meltdown at Three Mile Island, Pennsylvania, and Chernobyl. The students are asked to research some of these plants in order to discover what lead to the accidents that occurred there.

Unit two: Re-enactment of the Kyoto Protocol.

The beginning of the unit will involve a PowerPoint presentation and possible video that objectively looks at the topic of greenhouse gases to provide the science behind the controversy.

The second assignment for the class is to have pairs of students research a particular country that participated in the Kyoto conference, and prepare a short (five minute) PowerPoint presentation explaining their country's position. Each of the presentations is designed to gather a worldwide perspective on global warming without debate. The students will be responsible for picking their country from a list supplied by the teacher, and producing their own research using a variety of sources.

The third activity involves a role-playing activity based on the Kyoto Protocol. This activity was developed by the National Center for Case Study Teaching in Science (2006). Four student groups are formed (although this may be expanded, based on number of students in class), with a majority of the groups representing industrialized

Appendix E: Continued

countries, and a minority of the groups representing developing countries. Each industrialized country starts with a number of carbon dioxide “units” and money, with the overall goal of decreasing the number of carbon dioxide units as a class. Scoring is based on the combination of carbon dioxide units and money. Developing countries get scored based on the number of carbon dioxide units sold to the industrialized countries. This activity will provide the students an understanding of the negotiations that take place between groups in order to reach a consensus.

Unit three: Stem Cells: Saving Superman (National Center for Case Study Teaching in Science, 2003).

Part I of the unit involves three readings with concept check questions following each section. The first short reading involves the basic biology behind stem cells, including what stem cells are and the three types of stem cells. The second reading involves the harvesting of stem cells, including through embryonic tissue and fetal tissue, the main point of contention for many people. The third reading involves the possible applications of stem cell research, including the use of adult stem cells and potential of embryonic stem cells.

Part II involves a “Role Play/Jigsaw” involving six groups with views on stem cell research: stem cell researchers, pharmacologists, senators, ethicists from the National Bioethics Advisory Commission (NBAC), right to life members, and patients with autoimmune diseases and other disorders. The first phase involves each of the groups researching their position on the topic. Senators should research current laws and guidelines regarding stem cell research. Following a research period, the class

Appendix E: Continued

redistributes into new groups (jigsaw), with one member from each perspective forming a group, with the senator as facilitator. Following the small group debate, the senator is responsible for crafting a final position for the group and presenting the position to the class.

Part III is not part of the original case study, however, a whole class discussion regarding students' actual views and subsequent individual writing assignment (two to three pages) will allow individuals to express themselves without being confined to a role. The individual papers also provide data regarding emergent argumentation patterns throughout the course of the study, rather at a beginning and end point of the semester.

Unit four: Transgenic Crops: Do You Really Know What You're Eating?

The first phase of the case is to provide two competing briefs regarding the safety of genetically modified foods, and to provide scientific background for the case. In this story, the parents investigate a civil action against many of the groups potentially responsible for the boy's death, including the Environmental Protection Agency, BioCrystal (a fictitious company who produced the genetically modified seeds), Taco Heaven (the fast food restaurant), Stacey Brands (the company that produced the genetically modified foods), and the farmers themselves.

Following the briefs, there are a series of scientific questions regarding genetic engineering and genetically modified foods. Each group of students is to pick one of the questions, research it, and prepare a ten minute PowerPoint presentation to the class. This information becomes the content knowledge needed to participate in the issue part of the case.

Appendix E: Continued

Following the content presentations, the students are divided into roles to perform a mock trial regarding the civil case. Students will act as lawyers, spokespersons for the defendants, jury, and expert witnesses. Each student will serve as an expert witness regarding the question he or she presented in the previous step. Following the trial, the class will perform a “fishbowl discussion,” in which the jury first deliberates the extent to which each of the defendants were responsible for the tragedy, and then the rest of the class may weigh in with their own ideas about culpability.

The last assignment is for each student to write a letter regarding genetically modified foods to a U.S. Senator or Representative with their position on genetically modified foods. The letter can be written from either a pro or con perspective, but must be persuasive and incorporate scientific knowledge in the letter. This assignment will be between 250-500 words long.

Unit Five: Pesticides: Can We Do Without Them?

Part I of the unit provides the background and sets the stage for the debate. Although there is little content involved, the students are asked to identify multiple stakeholders as well as provide information about pests and pesticides. This section will provide the student with some of the background, as well as ask the student questions regarding the underlying themes of SSI, including the ability to make ethical decisions and participate from the viewpoint of multiple stakeholders.

Part II of the unit provides some of the reasoning behind each of the stakeholders in the debate. I would extend this section of the unit by splitting each of the classes into the four stakeholder groups mentioned in Part I, and have them research and debate the

Appendix E: Continued

issue in a whole class format. This would provide the students with much more information than provided by the unit, and also involve the students with direct debate experience. At the end of the debate, the students will be able to answer the questions listed at the section, which include the benefits and potential dangers of pesticides raised during the debate.

Part III involves the vote by the county commissioner. In this case I would have each of the students play the role of county commissioner to cast a vote regarding pesticide use. There are four questions at the end of this section asking the students to consider the political, social, and economic issues of the pesticide ban, as well as the ethics of his decision. These questions serve to complete the debate, and will provide

Appendix F: Persuasive Essay Assignment

Please take a few minutes to read the following story,

Cystic Fibrosis (CF) is an autosomal recessive genetic trait. It is one of the most prevalent genetic diseases. In England and the United States one out of 2,000 newborns is affected and one out of 20 people is a carrier.

Cystic fibrosis causes a deficient functioning of the external secretion glands that is pronounced (among other things) in the production of salty sweat, in digestion disorders, and in the production of large quantities of mucus in the respiratory tracts. The mucus causes recurrent lung infections. Each additional infection adds to the long-term damage to the lungs. The disease is therefore lethal: patients rarely survive past the age of 40.

The gene responsible for CF has been located. Scientists from laboratories in several countries are now working on methods for genetic therapy. One idea was to substitute a healthy gene for the deformed one in the lung tissue. However, the complex branching of the lungs makes it impossible to remove the Epithelium cells and then return them after the gene substitution. In 1992 one group of researchers succeeded in inserting the gene into the Epithelium of a rat's lung where it continued to function for 6 weeks. Another research direction focuses on the development of a spray consisting of normal genes attached to transporters whose role will be to insert the genes into the cells. The idea is that patients will inhale the spray from time to time (with the hope that the normal genes will be able to function in the cells). Despite all these efforts, it is still a long way before genetic treatment of CF can become practical. Meanwhile, patients keep suffering.

Appendix G: Persuasive Essay Assignment (posttest 2)

The Fluoridation Debate

For over 50 years, cities and towns have been adding fluoride to their water supply. Supporters of this practice suggest that adding fluoride to the public water supply is beneficial for teeth. Opponents argue that fluoride should not be added to water supplies because it forces citizens to consume a chemical that is not necessary and may be hazardous. Below you will find statements which describe both sides of the argument.

Fluoridation Is Beneficial

Adding fluoride to the public water supply is a simple, cheap, and effective way to promote dental health. It works to strengthen tooth enamel, which is the outer layer of a tooth and which forms a primary protection against cavities. The presence of fluoride in the water has been shown in scientific studies to drastically reduce tooth decay. The occurrence of cavities in babies and children can be reduced by as much as 60% by drinking water with adequate amounts of fluoride, and cavities can be reduced by as much as 35% in adult teeth. The dental health of a community can be greatly improved by water fluoridation. Currently, 360 million people worldwide and 145 million in the United States benefit from water fluoridation.

Cities and towns, which do add fluoride to their water supplies, monitor the amount of fluoride which exists naturally in the water. Based on this information, they add just enough to improve the dental health of their citizens. The typical concentration of fluoride in a fluoridated water supply is 1 part per million. (For every million water molecules, there is only one fluoride molecule).

Fluoridation is Harmful

Adding fluoride to the public water supply can be dangerous and should be stopped. Fluoride is a toxic substance that has been linked to diseases such as cancer and fluorosis. Dental fluorosis is caused by the accumulation of too much fluoride in teeth. This condition can cause pits and marks in teeth. The number of individuals affected by dental fluorosis has been shown to increase from 15-65% in cities and towns which fluoridate their water supplies. Dental fluorosis can also lead to skeletal fluorosis. Whereas dental fluorosis may only affect a person's appearance, skeletal fluorosis is far more serious. It begins by causing symptoms similar to arthritis and can end up crippling sufferers.

The scientific community has known about the toxic effects of fluoride for many years. In fact, it was used throughout World War I as rat poison. A substance strong enough to kill rodents should not be added to water supplies used by millions of people. Even though the amount of fluoride added to water supplies may be low, the chemical can accumulate in individuals and eventually cause serious health problems.

The addition of fluoride to a public water supply forces all of the users of that water to consume the substance. This practice amounts to forced medication. Forcing individuals to take substances without their direct consent, as in this case, is unethical and should not be permitted. Government officials should not have to right to force people to take substances they may not wish to have. But that is exactly what happens when fluoride is added to public water supplies.

Appendix H: Rubric for analysis of written argumentation (WAR)

Criterion	Score PRE	Score POST	Description
Justifications (Zohar and Nemet, 2002)	2	2	Two or more valid justifications
	1	1	One valid justification
	0	0	No justifications offered
Structure (Zohar and Nemet, 2002)	2	2	A complex structure with justification supported by another reason.
	1	1	A simple structure consisting of a conclusion supported by at least one reason
	0	0	No valid justification
Subject Matter Knowledge (Walker and Zeidler, 2007)	3	3	Correct consideration of specific evidence claims or SMK.
	2	2	Consideration of non-specific evidence claims or SMK.
	1	1	Incorrect consideration of evidence claims or SMK.
	0	0	No evidence claims or subject matter knowledge (SMK) are considered.

TEST	J	S	SMK	TOTAL
PRE	_____	_____	_____	_____
POST	_____	_____	_____	_____

Appendix I: Argumentation Interview

Please take a few minutes to read the following story,

Cystic Fibrosis (CF) is an autosomal recessive genetic trait. It is one of the most prevalent genetic diseases. In England and the United States one out of 2,000 newborns is affected and one out of 20 people is a carrier.

Cystic fibrosis causes a deficient functioning of the external secretion glands that is pronounced (among other things) in the production of salty sweat, in digestion disorders, and in the production of large quantities of mucus in the respiratory tracts. The mucus causes recurrent lung infections. Each additional infection adds to the long-term damage to the lungs. The disease is therefore lethal: patients rarely survive past the age of 40.

The gene responsible for CF has been located. Scientists from laboratories in several countries are now working on methods for genetic therapy. One idea was to substitute a healthy gene for the deformed one in the lung tissue. However, the complex branching of the lungs makes it impossible to remove the Epithelium cells and then return them after the gene substitution. In 1992 one group of researchers succeeded in inserting the gene into the Epithelium of a rat's lung where it continued to function for 6 weeks. Another research direction focuses on the development of a spray consisting of normal genes attached to transporters whose role will be to insert the genes into the cells. The idea is that patients will inhale the spray from time to time (with the hope that the normal genes will be able to function in the cells). Despite all these efforts, it is still a long way before genetic treatment of CF can become practical. Meanwhile, patients keep suffering.

Interview questions

1. Rebecca and Joseph both have brothers who are sick with CF (this means that each of them MAY carry the gene for CF). Rebecca and Joseph got married and Rebecca is now pregnant. Should they abort the embryo (which MAY have CF)? Explain?
2. Genetic tests showed that Rebecca and Joseph are carriers of CF and that the embryo is homozygous for CF. Rebecca and Joseph contemplate whether or not they should have an abortion.
 - a. What is the moral problem under consideration?
 - b. Do you think they should perform an abortion? Offer reasons for your position?
 - c. Your friend disagrees with you. Define his or her position. Offer reasons for that position (what will your friend say to convince you that s/he is right)?
 - d. How would you answer your friend? Explain!

Appendix I: Continued

A dilemma taken from everyday life

Sue, a ninth grade teacher, became fed up with students who are cheating on tests. She decided to offer her students a new system of *Honor Tests*. Until now, tests in her classroom took place in the common manner in which the teacher guards fiercely against students who are trying to cheat. According to her suggestion, she would leave the classroom during tests and students would promise not to cheat.

Sue presented her new idea to her students. Following a class discussion, all students consented to try the Honor Tests system. It was agreed that by the end of each test, the teacher will ask whether students indeed kept their promise. Students continued to discuss this issue during recess. The general opinion among students was that such a test shows a high level of confidence between a teacher and her students, and therefore, if the teacher is willing to trust her students they should keep their word and not cheat.

Two weeks later Sue gave a History test and walked out of the classroom. During the test, Rick noticed that Andy, his best friend, was copying from some notes he prepared in advance. Rick was disappointed with his friend because he felt that his behavior undermined the efforts of everybody else in the class. Rick tried to convince Andy to confess, but Andy was not willing to do so.

In class the following day, Sue asked her students whether or not they had kept their promise.

Interview questions

1. What is the moral problem under consideration?
2. Do you think Rick should tell the teacher that Andy cheated? Offer reasons for your position!
3. Your friend disagrees with you. Define his/her position. Offer reasons for that position (what will your friend say to convince you that s/he is right?)
4. What will you answer your friend? Explain!

ABOUT THE AUTHOR

Brendan Callahan received Bachelor's Degrees in Biology (1997) and Chemistry (2001), and a M.Ed. in Curriculum and Instruction (2004) from the University of South Florida. He taught science and math at a private school prior to his enrolling in the Master's program, then continued to teach full time as a middle and high school science teacher while as a graduate student.

As a Ph.D. student at the University of South Florida, Mr. Callahan was actively involved in teaching at the college level and research. He had the opportunity to teach undergraduate methods classes for the College of Education, and he presented at national conferences in science education. He also coauthored a paper in 2009 that appeared in the *Journal of Research in Science Teaching*.