2011

Using Argument-Driven Inquiry to Enhance Students' Argument Sophistication When Supporting a Stance in the Context of Socioscientific Issues

Jonathon A. Grooms
USING ARGUMENT-DRIVEN INQUIRY TO ENHANCE STUDENTS’ ARGUMENT
SOPHISTICATION WHEN SUPPORTING A STANCE IN THE CONTEXT OF
SOCIOSCIENTIFIC ISSUES

By

JONATHON A. GROOMS

A Dissertation submitted to the
School of Teacher Education
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Spring Semester, 2011
The members of the committee approve the dissertation of Jonathon Grooms defended on March 16, 2011.

__________________________________
Victor Sampson
Professor Directing Dissertation

__________________________________
Simon Capstick
University Representative

__________________________________
Sherry Southerland
Committee Member

__________________________________
Allan Jeong
Committee Member

Approved:

__________________________________
Lawrence Scharmann, Chair, School of Teacher Education

The Graduate School has verified and approved the above-named committee members.
I dedicate this dissertation to my mother, Linda Grooms and my late father, Nolie Grooms, Jr. for fostering my natural curiosity, instilling a love of learning at an early age, and for teaching my sisters and me the importance of an education. The fact that my sisters and I are all active members of the education community, following in the footsteps of our parents, speaks volumes to the influence they have had on our lives. Through their example I have learned the value of setting lofty goals and now experience the gratification of attaining them.

I also dedicate this dissertation to my wife, Lorie, for supporting me through this process. Her encouragement, despite long hours and time apart, has served as strong motivation. Her selflessness in allowing me to pursue a challenging personal goal, while pushing me ever closer, is truly appreciated.
ACKNOWLEDGEMENTS

I would like to acknowledge the members of my doctoral committee, Drs. Sherry Southerland, Allan Jeong, and Simon Capstick, for their guidance and leadership during the course of this research project and throughout my graduate studies. The mentorship I have received from my major professor, Dr. Victor Sampson, has shown me what it means to be a quality researcher and professor and for that I am thankful. With regard to this particular research project I would like to thank Dr. Stephanie Dillon and Dr. Joi Walker for facilitating access to the student participants who participated in this study, without whom, this project would not have been possible. Special thanks are due to my colleagues in The Florida State University’s Office of Science Teaching Activities, for their support during my graduate studies and availability and willingness to engage in productive conversations surrounding my research. Thank you to Barry Golden, a colleague and classmate, for taking the time to assist in scoring 600-plus arguments for this research project, while in the midst of his own dissertation research. Finally, thank you to my fellow classmates, particularly Lance King, Joi Walker, Patrick Enderle, Barry Golden, and Melanie Hester for challenging my thinking throughout my graduate studies and the course of this research project. Their valuable perspectives and critiques served to enhance my graduate experience and the quality of this research project.
TABLE OF CONTENTS

Introduction ................................................................................................................................. 1
Theoretical Foundations ............................................................................................................... 2
  Reasons for integrating argumentation into science education ........................................... 2
  Factors influencing how students participate in scientific argumentation ................................ 5
  Achieving scientific literacy ................................................................................................. 8
Research Questions .................................................................................................................. 10
Research Design ....................................................................................................................... 12
  Participants ............................................................................................................................ 12
  Instructional context ........................................................................................................... 13
    Argument-driven inquiry (ADI) ........................................................................................... 14
    Comparison of instructional contexts ............................................................................... 15
  Procedures ............................................................................................................................ 17
  SSI Tasks ............................................................................................................................. 18
Data Collection Instruments ................................................................................................... 21
  Science content knowledge assessment ............................................................................ 21
  Epistemological sophistication questionnaire .................................................................... 22
  SSI follow-up questionnaire .............................................................................................. 25
Data Analysis ........................................................................................................................... 26
  Science content knowledge ............................................................................................... 26
  Epistemological sophistication ............................................................................................ 27
  Argument quality ................................................................................................................ 28
    Argument structure ........................................................................................................... 28
    Nature of the argument justifications .............................................................................. 30
Results and Discussion ............................................................................................................. 32
  Influence of various factors on students’ votes related to SSI .............................................. 32
  Nature of justifications supporting students’ votes on SSI ................................................... 35
    SSI specific justifications ................................................................................................. 38
    Differences in justifications between groups .................................................................. 43
    Post-intervention distribution of justifications and influential factors .............................. 46
  Structure of arguments supporting students’ votes on SSI ................................................... 49
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science Content Questions related to the EPA task</td>
<td>91</td>
</tr>
<tr>
<td>Science Content Questions related to the Sugar Task</td>
<td>92</td>
</tr>
<tr>
<td>Epistemological Sophistication Assessment</td>
<td>94</td>
</tr>
<tr>
<td>Instructions</td>
<td>94</td>
</tr>
<tr>
<td>EPA Task Follow-up Questionnaire</td>
<td>96</td>
</tr>
<tr>
<td>Instructions</td>
<td>96</td>
</tr>
<tr>
<td>Sugar Task Follow-up Questionnaire</td>
<td>98</td>
</tr>
<tr>
<td>Instructions</td>
<td>98</td>
</tr>
<tr>
<td>Institutional Review Board Approval Letter</td>
<td>100</td>
</tr>
<tr>
<td>Approved IRB Consent Form</td>
<td>102</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1.1 Levels of epistemological understanding (Kuhn, D. et al., 2000, p. 311, original emphasis) ................................................................................................................................ 7

Table 2.1 Demographics for comparison and treatment groups and institutions .................. 13

Table 2.2 Course objectives, student outcomes for each group .............................................. 14

Table 2.3 Phases of the ADI instructional model (Sampson, Walker, and Grooms, 2009, p. 43) 15

Table 2.4 Features of each instructional model ........................................................................... 16

Table 2.5 Distribution of participants by epistemological category across three judgment domains in Kuhn, et al. (2000) and Grooms pilot studies ................................................................. 24

Table 2.6 Percentage of participants categorized as evaluativist on three different judgment domains (Kuhn, D. et al., 2000) ....................................................................................................................... 25

Table 2.7 Coding scheme for epistemological sophistication within sub-scale ....................... 28

Table 2.8 Levels of argument structure ...................................................................................... 29

Table 2.9 Types of argument justifications and descriptions ...................................................... 31

Table 3.1 Mean rating for factors influencing students’ votes, pre-intervention ......................... 33

Table 3.2 Mean rating for factors influencing students’ votes, by group, pre-intervention ........ 34

Table 3.3 Mean rating for factors influencing students’ votes, by task, pre-intervention .......... 35

Table 3.4 Number of justifications and average justifications for each group on each task, pre-intervention .......................................................................................................................... 36

Table 3.5 Mean rating for factors influencing students’ votes, by task, post-intervention ......... 48
LIST OF FIGURES

Figure 3.1 Overall distribution of pre-intervention justifications for both tasks and both groups 37
Figure 3.2 Distribution of justifications for each SSI task, pre-intervention ............................ 39
Figure 3.3 Comparison of distribution of justifications for the EPA and Sugar tasks between groups, pre-intervention ................................................................. 45
Figure 3.4 Comparison of distribution of justifications for the EPA and Sugar tasks between groups, post-intervention ................................................................. 47
Figure 3.5 Distribution of argument justifications for both groups on each task, post-intervention ........................................................................................................ 48
Figure 3.6 Distribution of argument scores for both groups, pre-intervention ............................ 51
Figure 3.7 Distribution of argument scores for each SSI task, pre-intervention .......................... 54
Figure 3.8 Distribution of argument scores for each SSI task, post-intervention ...................... 56
Figure 3.9 Epistemological sophistication for the comparison and treatment groups, pre- and post-intervention .......................................................................................... 60
Figure 3.10 Distribution of argument scores by different levels of epistemological sophistication for all participants, pre-intervention .......................................................... 62
Figure 3.11 Distribution of argument scores for students classified as multiplist, pre- and post-intervention .............................................................................................. 64
Figure 3.12 Distribution of argument scores for students classified as evaluativist, pre- and post-intervention .............................................................................................. 65
Figure 3.13 Distribution of argument scores by content knowledge for both groups, pre-intervention ............................................................................................................ 68
Figure 3.14 Distribution of argument scores by content knowledge for both groups, post-intervention ............................................................................................................. 69
Figure 3.15 Distribution of argument scores for students with low content knowledge by group, post-intervention ...................................................................................... 71
ABSTRACT

This quasi-experimental study assesses the extent to which the Argument-Driven Inquiry (ADI) instructional model enhances undergraduate students’ abilities to generate quality arguments supporting their stance in the context of a Socioscientific Issue (SSI) as compared to students experiencing a traditional style of instruction. Enhancing the quality of undergraduate students’ arguments in the context of SSI can serve as an indirect measure of their scientific literacy and their ability to make sound decisions on issues that are inherently scientific but also involve social implications. Data collected in this study suggest that the undergraduate students experiencing the ADI instruction more readily provide rationales in their arguments supporting their decisions regarding two SSI-tasks as compared to a group of undergraduate students experiencing traditional instruction. This improvement in argument quality and gain in scientific literacy was achieved despite the overall lower SSI related content knowledge of the ADI students. Furthermore, the gap between the argument quality of those students with high versus low SSI related content knowledge was closed within the ADI group, while the same gap persisted post-intervention within the traditional instruction students. The role of students’ epistemological sophistication was also investigated, which showed that neither instructional strategy was effective at shifting students’ epistemological sophistication toward an evaluativist stance. However, the multiplists within the ADI group were able to significantly increase the sophistication of their arguments whereas the traditional students were not. There were no differences between the quality of arguments generated by the evaluativist students with either the treatment or comparison groups. Finally, the nature of the justifications used by the students revealed that the students (both comparison and treatment groups) did not invoke science-based justifications when supporting their stance, despite students’ self-reports that scientific content knowledge accounted for the greatest influence on their stance, related to the SSI tasks. The results of this study suggest that the scientific habits of mind the students learned in the context of ADI investigations are transferred to the novel SSI contexts. Implications for the use of argument-based instructional models to enhance the generation of socioscientific arguments and to promote the development of scientific literacy are also discussed.
Science literacy is an important goal for science education in both the K-12 and university systems (AAAS, 1989; Halyard, 1993; NRC, 1996a, 1996b; NSF, 1998; NRC, 1999, 2007). An individual that has achieved scientific literacy understands scientific content, processes, and the epistemological commitments of science, such that one can invoke these understandings when making decisions in the context of relevant social issues (NRC, 1996a). Given the significance of scientific literacy, it is important for teachers to choose and implement instructional strategies that are designed to help students achieve this goal.

Instructional strategies centered on the practice of scientific argumentation, which require the generation, evaluation, critique and refinement of data and evidence in an attempt to explain natural phenomena (Duschl and Osborne, 2002; Osborne, Erduran and Simon, 2004; Sampson and Grooms, 2009a; Sampson, Grooms and Walker, 2009b), can help students develop and use scientific habits of mind (Yerrick, 2000; Sandoval and Reiser, 2004) and understand the content of science (Bell and Linn, 2000; Zohar and Nemet, 2002). Scientific argumentation is also a key epistemic activity of the scientific community. Thus, instructional strategies grounded in scientific argumentation can also give students an opportunity to learn how to participate in the process of science, which is a key step in the development of scientific literacy (Driver, Newton and Osborne, 2000; Duschl and Osborne, 2002; Osborne et al., 2004; Jimenez-Aleixandre, 2008). Furthermore, the ability to generate, evaluate, critique and refine data and evidence are foundational skills necessary for engaging in complex discussions related to socioscientific issues (Driver et al., 2000; Kolsto, Bungum, Arnesen et al., 2006; Sadler and Fowler, 2006b; Sadler, Barab and Scott, 2007). Thus, the use of scientific argumentation as an instructional strategy seems appropriate as a means to help students develop the ability to use the content, processes, and epistemological commitments of science in order to make decisions about relevant social issues.
The primary goal of this study is to explore the effect of *Argument-Driven Inquiry* (ADI) (Sampson et al., 2009b) in supporting students’ abilities to produce quality arguments supporting their stance on a socioscientific issue. However, there are myriad factors that can influence how students engage in scientific argumentation, such as content knowledge (Sadler and Fowler, 2006b; von Aufschnaiter, Erduran, Osborne et al., 2008), interest in the topic (Dole and Sinatra, 1998; Nussbaum and Bendixen, 2003a), ability to argue or understanding of arguments and argumentation (Kuhn and Udell, 2003; Nussbaum and Bendixen, 2003a; Nussbaum and Sinatra, 2003b), and an individual’s personal epistemology or how they view knowledge and learning (Kuhn, 1993; Nussbaum, Sinatra and Poliquin, 2008). Therefore, in addition to analyzing students’ arguments this study will also examine the how students’ background knowledge and epistemological sophistication influence their ability to generate quality arguments. Finally, given the potential for such instructional strategies to influence students’ epistemological sophistication (Kuhn, Cheney and Weinstock, 2000), a final goal of this study is to measure the extent to which students’ epistemological sophistication changes after experiencing a series of ADI activities.

**Theoretical Foundations**

*Reasons for integrating argumentation into science education*

Teaching for scientific literacy requires the coordination of the epistemological, procedural and conceptual basis of science, which allows us to discuss “how we know what we know and why we believe it” (Duschl, 2008, p. 269). Pedagogy that bridges epistemological, procedural, and conceptual aspects of science is largely missing from the science classroom in that conceptual and epistemic learning are not situated in curricula, instruction, and assessment such that both are happening concurrently (Duschl, 2008). One strategy that has been proposed by science educators as a way to integrate epistemic, procedural, and conceptual factors is through the use of scientific argumentation as an instructional model to aid the teaching and learning of science. The process of scientific argumentation is critical to the advancement of sound scientific knowledge claims because it involves critiquing the conceptual products, the methods
used for collecting data, and the analytical procedures used to coordinate the data and subsequent knowledge claim (Driver et al., 2000; Duschl and Osborne, 2002). Critiques based on these criteria move the epistemological, conceptual, procedural, as well as social aspects of scientific knowledge development to the forefront of classroom activity, thus moving science education “from a dominant focus on conceptual learning toward a more balanced focus among things conceptual, epistemic, and social” (Duschl, 2008, p. 283).

Implementing scientific argumentation into the teaching and learning of science, however, can be difficult in a classroom setting. In order to engage students in argumentation, teachers must restructure classroom norms from traditional pedagogical approaches (e.g. lectures or ‘cookbook’ labs) to those that allow for the production and critical evaluation of scientific arguments in an attempt to make sense of a phenomenon (Osborne et al., 2004). To help accomplish this transformation, science educators have designed curricula (Krajcik, McNeill and Reiser, 2008), technology-enhanced learning environments (Bell and Linn, 2000; Suthers and Hundhausen, 2003; Sandoval and Reiser, 2004), and instructional models (Sampson and Grooms, 2009a; Sampson et al., 2009b) that teachers can use to promote and support scientific argumentation inside the classroom. These approaches, in general, are designed to give students more opportunities to construct and critique explanations or arguments of a natural phenomenon as part of the inquiry process. Empirical research suggests that these approaches not only improve students’ abilities to develop and critique arguments, but also enhance conceptual understanding (Bell and Linn, 2000; Zohar and Nemet, 2002; Sandoval and Millwood, 2005; Nussbaum et al., 2008). Yet, few researchers have attempted to determine if these methods encourage people to use what they learned inside the science classroom in the context of a relevant social issue.

Another approach that is often used to help students develop scientific literacy is to encourage students to explore, develop, and debate possible solutions to socioscientific issues (SSI) (Sadler, 2004a); for example, discussing genetic dilemmas (e.g. genetic engineering, applied human genetics) during a biology course (Zohar and Nemet, 2002) or conducting investigations of water pollution and potential actions for remedying the problem (Sadler et al., 2007). SSI such as these are complex, ill-structured problems that
do not have a straight-forward solution and often require the coordination of various perspectives, such as moral, political, social, and economic concerns (Sadler, 2004a).

This line of research suggests that when students engage in learning tasks that involve SSI they can develop many of the same skills that are important for scientific argumentation, namely, evaluation of evidence, construction of arguments, and evaluation of competing arguments. SSI-based instruction can also enhance students’ content knowledge (Zohar and Nemet, 2002; Sadler, 2004a, 2009b). However, many science educators do not use this type of approach inside the classroom, potentially because of the amount of time required to complete such activities (Sadler et al., 2007) or perhaps based on the misconception that SSI curricula lack an emphasis on science content (Roberts, 2007).

Determining which instructional method to use in a particular course involves weighing the various outcomes the course is designed to foster in students. Take, for instance, an introductory undergraduate science course, which is often divided into a lecture and lab section. The goal of the lecture is to present a body of knowledge to students, often based on a transmission model of learning (Wyckoff, 2001; DeHaan, 2005), whereas the lab is designed to illustrate these concepts, or simply confirm them, and (perhaps) to help students develop inquiry skills (Cooper and Kerns, 2006). It can be argued that these introductory courses are the most important component of science, technology, engineering, and math programs because of the sheer number and diversity of students that enroll in them and with the consideration that these are the only science courses many undergraduate students experience (Halyard, 1993; Rothman and Narum, 1999). Meeting the needs of the future scientists, engineers, teachers, and business leaders that may all be enrolled in the same course is no easy task. A heavy emphasis on science content may be most appropriate for a sub-group of students in these courses while a heavy emphasis on social implications may be more appropriate for those not strongly invested in learning science content. Perhaps it is possible to address the needs of these various sub-groups while still promoting scientific literacy by implementing argumentation based pedagogies inside the lab.

One particular instructional strategy with the potential to promote scientific literacy in the undergraduate lab is Argument-Driven Inquiry (ADI) (Sampson et al.,
2009b; Walker, Grooms, Anderson et al., 2010, March). The ADI instructional model affords students the opportunity to generate data and evidence in order to address an overarching research question or as an attempt to explain a natural phenomenon. The students then use their data and evidence to form a scientific argument that is presented to and critiqued by their classmates, thus providing educative feedback for the further refinement of the groups’ arguments. The ADI instructional model is intended to help students develop inquiry skills (e.g., how to gather and analyze data), reasoning skills (e.g., understanding how to support claims with evidence), and important scientific habits of mind (e.g., understanding the value and limitations of data) (Walker et al., 2010, March).

In a recent study involving the ADI instructional model, students who participated in a semester-long course implementing ADI produced better quality arguments explaining natural phenomena than their peers in a non-ADI course. Despite conducting fewer laboratory activities students in the ADI approach showed no difference in their science content knowledge (Walker et al., 2010, March). Considering the various aspects of scientific literacy, there is evidence to suggest that the students in the ADI course developed an equivalent understanding of the material and were better able to craft scientific arguments; however, their ability to transfer these argumentation abilities about natural phenomena to a socioscientific context (an important indicator of scientific literacy) has not been examined.

**Factors influencing how students participate in scientific argumentation**

As mentioned above, there are a variety of factors that may help or hinder how students participate in scientific argumentation when they are encouraged to engage in this complex practice. Although content knowledge and interest may be addressed through the careful selection of topics and appropriate scaffolding for argumentation-based activities, accounting for variation in students’ epistemological commitments (e.g., how an individual views knowledge and its construction; either as certain, tentative, socially constructed, etc.) is much more difficult. Furthermore, in a study by Mason and Scirica (2006), it was found that a student’s level of epistemological sophistication was a
more significant predictor of argument quality than either content knowledge or interest in the subject.

Kuhn et al. (2000) provides a useful model of epistemological development, which can be used to describe the epistemological sophistication of students. Kuhn and her colleagues suggest that people move from an absolutist conception of knowledge to an evaluativist conception of knowledge over time, with the multiplist conception somewhere in between. During this progression, it is proposed that a person’s view of knowledge moves from a conception that knowledge is objective and external from the knower (absolutist stance) to the conception that knowledge is subjective or uncertain, however these uncertain knowledge claims can be evaluated (evaluativist stance). The evaluativist differs from the multiplist – whom also views knowledge as subjective and uncertain – in that the multiplist views all knowledge claims as equal (Hofer and Pintrich, 1997; Kuhn et al., 2000). Table 1.1 identifies the various levels of epistemological understanding associated with this framework.

If a goal in science education is for students to be able to produce high-quality scientific arguments we must also concern ourselves with helping them develop their epistemological sophistication because such development is related to their learning and interpretation of complex information (Schommer, 1990, 1994). It has been hypothesized that epistemological development (Bendixen and Rule, 2004) proceeds from an absolutist through multiplist, ultimately to an evaluativist stance (Kuhn et al., 2000) via a possible mechanism similar to cognitive restructuring models associated with the process of conceptual change (Dole and Sinatra, 1998; Bendixen and Rule, 2004). If that is the case, then perhaps, argumentation-based pedagogies may be useful for engendering the cognitive dissonance needed for this development to occur.

For example, consider a classroom environment that promotes and supports scientific argumentation. A student in such a classroom whom has a less-developed personal epistemology (e.g. that knowledge is absolute, handed down from authority, and that learning is quick and easy) could encounter some cognitive dissonance when their personal set of epistemological commitments were called into question by the nature of the classroom instruction. In a classroom environment that values scientific argumentation, e.g., a classroom structured around argument-driven inquiry, students are
confronted with situations where an absolutist personal epistemology is no longer fruitful or well aligned with the nature of classroom instruction (e.g., students must develop their own knowledge, knowledge is critiqued by peers and the development of knowledge takes time and refinement). It is possible that a conflict between personal commitments and those valued in the classroom environment may initiate in the student a reevaluation of their personal epistemological commitments or result in their adoption of new commitments and ultimately progress toward a more sophisticated personal epistemological stance.

Table 1.1 *Levels of epistemological understanding* (Kuhn et al., 2000, p. 311, original emphasis)

<table>
<thead>
<tr>
<th>Level</th>
<th>Assertions</th>
<th>Reality</th>
<th>Knowledge</th>
<th>Critical Thinking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolutist</td>
<td>Assertions are FACTS that are correct or incorrect in their representation of reality (possibility of false belief).</td>
<td>Reality is directly knowable.</td>
<td>Knowledge comes from an external source and is certain.</td>
<td>Critical thinking is a vehicle for comparing assertions to reality and determining their truth or falsehood.</td>
</tr>
<tr>
<td>Multiplist</td>
<td>Assertions are OPINIONS freely chosen by and accountable only to their owners.</td>
<td>Reality is not directly knowable.</td>
<td>Knowledge is generated by human minds and is uncertain.</td>
<td>Critical thinking is irrelevant.</td>
</tr>
<tr>
<td>Evaluativist</td>
<td>Assertions are JUDGEMENTS that can be evaluated and compared according to criteria of argument and evidence.</td>
<td>Reality is not directly knowable.</td>
<td>Knowledge is generated by human minds and is uncertain.</td>
<td>Critical thinking is valued as a vehicle that promotes sound assertions and enhances understanding.</td>
</tr>
</tbody>
</table>

Consider the importance of scientific argumentation in the context of socioscientific issues with the underlying factor of personal epistemology. Given the
variety of perspectives that an individual may encounter when dealing with a SSI, it is important for that individual to evaluate the merits of the competing arguments. Each argument is likely well reasoned, however, they may not be equally defensible (Sadler et al., 2007). An individual with a personal epistemology consistent with an evaluativist perspective will be inclined to evaluate the competing knowledge claims or arguments in an attempt to determine and judge the soundness of the various claims (Kuhn et al., 2000) when making a decision relative to the SSI. Whereas an individual operating from a multiplist perspective would view all competing arguments as equally valid or perhaps worse yet, an absolutist may unquestioningly accept the argument of anyone perceived to be an expert. It is my contention that the scientific habits of mind that are promoted in argument-driven inquiry activities, such as the critical evaluation of evidence and the critique of counterarguments, will serve as a catalyst for shifting students’ epistemological commitments toward those associated with an evaluativist stance.

Achieving scientific literacy

In order to realize the goal of scientific literacy students need to develop a conceptual understanding of scientific content and processes as well as be able to apply those understandings to relevant social issues (AAAS, 1989; Halyard, 1993; NRC, 1996a). “Modes of argumentation and the epistemology of science are an important component of any education that seeks to enhance scientific literacy” (Driver et al., 2000, p. 305). Argumentation in the context of socioscientific issues is also valuable for promoting scientific literacy (Osborne et al., 2004; Roberts, 2007; Sadler et al., 2007); however, for a variety of factors mentioned above, this strategy does not seem to be widely implemented in science classrooms, particularly at the college level. Furthermore, research shows that when individuals are dealing with SSI, they recognize the need to evaluate information and evidence, but they lack the skills to do so (Sadler, 2004a). This inability further strengthens the need for science instruction to emphasize the evaluation and critique of data and evidence that is the hallmark of argumentation-based instructional strategies, particularly argument-driven inquiry (Sampson et al., 2009b).

In the context of undergraduate science, the complete restructuring of large lecture courses to highlight argumentation and SSI to promote scientific literacy is likely
a tall order. However, it is more likely that the smaller lab portions of such courses could be reconceptualized to emphasize these components of teaching for scientific literacy. Furthermore, a shift from traditional laboratory approaches, described as “cookbook” labs, to one focused on inquiry and argumentation is potentially a smaller step – and thus more likely to happen – than a dramatic swing to a curriculum centered on SSI. Given the data presented above, argumentation-based instruction, and specifically argument-driven inquiry, has been successful in helping students develop better scientific arguments supporting explanations in regards to natural phenomena. This outcome may be sufficient for a laboratory course, but it does not engender the full scope of scientific literacy in that those argumentation skills are not used with respect to a relevant social issue.

“A major goal of schooling is to prepare students for flexible adaptation to new problems and settings. The ability of students to transfer provides an important index of learning that can help teachers evaluate and improve their instruction” (Donovan, Bransford and Pellegrino, 2000, p. 77), additionally, it is argued that “the amount of transfer will be a function of the overlap between the original domain of learning and the novel one” (Donovan et al., 2000, p. 63). Thus, in order to provide a realistic chance of having students transfer what they learn in the science classroom to situations they may encounter in their ‘everyday’ lives, the instructional strategies used in the classroom should provide opportunities to develop the skills and habits of mind needed for addressing new situations. ADI provides a context where scientific phenomena are the focus of investigation, yet the skills and scientific habits of mind associated with SSI-based tasks (i.e. evidence evaluation, construction of arguments, critique and evaluation of arguments) are central. The overlap of ADI and SSI activities does not represent a direct-transfer situation; however, ADI is positioned as a near-transfer situation that may prove beneficial to promoting students’ critical evaluation of socioscientific issues. Therefore, it is important to investigate whether or not students that develop robust skills in scientific argumentation, via argument-driven inquiry (Sampson et al., 2009b), are able to transfer those skills and habits of mind from the scientific context in which they were developed (i.e. in reference to natural phenomena) to a SSI-based context.

Finally, much research has been conducted to highlight the role of students’ epistemological commitments and the student’s ability to generate scientific arguments
(Kuhn, 1993; Kuhn et al., 2000; Sinatra, Southerland, McConaughy et al., 2003; Bendixen and Rule, 2004; Mason and Scirica, 2006; Nussbaum et al., 2008); however, there is scant research that speaks to the reflexive nature of this relationship. Although personal epistemology has been shown to influence learning (Schommer, 1994; Hofer and Pintrich, 1997; Bendixen and Rule, 2004) and how an individual engages in scientific argumentation (Mason and Scirica, 2006; Nussbaum et al., 2008), few studies have examined the extent to which learning to generate scientific arguments affects the development of an individual’s personal epistemology.

**Research Questions**

Given the discussion above I investigated the extent to which students participating in an *Argument-Driven Inquiry* (Sampson and Grooms, 2009a) laboratory course differed from students in a more traditional laboratory course in the quality of the arguments they were able to produce about their stance on a socioscientific issue. Additionally, given the important role of epistemological views for argument generation, evaluation and critique, as well as the potential for argumentation-based instruction to affect epistemological development, I investigated the extent to which students’ epistemological views shifted based on the nature of the course they experienced. Furthermore, I investigated the potential relationship between students’ epistemological views and the quality of their arguments about SSI and the potential relationship between students’ scientific content knowledge related to the SSI and the quality of their arguments.

Therefore, the following research questions guided this investigation:

1. Do students who completed a series of ADI lab activities craft stronger arguments about socioscientific issues, in terms of structure and nature of the justification, than students who completed a series of verification lab activities?

2. Do students who completed a series of ADI lab activities shift their epistemological views toward an evaluativist stance more than students who completed a series of verification lab activities?
3. Is there a relationship between students’ epistemological views, content knowledge, the nature of lab-instruction, and the quality of the arguments students produce when they are asked to justify a stance on a socioscientific issue?
CHAPTER TWO

Research Design

This study follows a quasi-experimental design comparing a treatment group of students experiencing classroom instruction aligned with the argument-driven inquiry model and a comparison group of students experiencing a more traditional style of instruction. Data collection occurred on a pre- & post-intervention time frame with the pre-intervention data collected at the beginning of the school semester and the post-intervention data collected just prior to the end of the semester, spanning fifteen weeks. The data collected consists of a science content knowledge assessment and a personal epistemology assessment. Additionally, the students completed two tasks, each presenting a different SSI, and then constructed a series of arguments supporting their stance about the issues presented in the tasks. The following sections will describe the participants, procedures, SSI tasks, and data collection instruments in more detail.

Participants

Two groups of students enrolled in an introductory chemistry laboratory course served as the participants for this research. The treatment group was taught using the Argument-Driven Inquiry instructional model while the comparison group experienced instruction more representative of a traditional laboratory approach. The students in the treatment group were enrolled in a large, 2-year, community college in the Southeastern U.S. The comparison group students were enrolled in the same course at a neighboring 4-year university. Six class sections at each institution participated in this study resulting in 73 student participants in the treatment group and 79 student participants in the comparison group. Approximately 84% of the treatment group and 75% of the comparison group identified themselves as science majors (or intending to major in science). The treatment group students were 57% female whereas the comparison group was slightly less with females accounting for 50% of the sample.

The two groups of students that participated in this study also reflect a representative sample of their respective institution in terms of the distribution of student
demographics. Table 2.1 shows the distribution of students by gender and ethnicity for both the treatment and comparison groups. Also included are the overall demographics for the home institution of each group. The data displayed in Table 2.1 indicate that the comparison and treatment groups had very similar compositions and that each group was reasonably representative of its home institution. Two possible exceptions are the proportions of white and black students within the treatment group with respect to their institution. The treatment group sample is composed of approximately 25% more white students than the distribution of their home institution as well as about 20% fewer black students than the distribution of their home institution. These differences may be a function of the particular class sections participating in this study, but may also be a result of students’ choosing to enroll in this particular chemistry laboratory course.

Table 2.1 *Demographics for comparison and treatment groups and institutions*

<table>
<thead>
<tr>
<th>Gender and Ethnicity</th>
<th>Comparison Group</th>
<th>Comparison Institution</th>
<th>Treatment Group</th>
<th>Treatment Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>50%</td>
<td>55%</td>
<td>57%</td>
<td>56%</td>
</tr>
<tr>
<td>Male</td>
<td>50%</td>
<td>45%</td>
<td>43%</td>
<td>44%</td>
</tr>
<tr>
<td>White</td>
<td>71%</td>
<td>68%</td>
<td>77%</td>
<td>51%</td>
</tr>
<tr>
<td>Black</td>
<td>13%</td>
<td>10%</td>
<td>14%</td>
<td>35%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>14%</td>
<td>12%</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>Asian</td>
<td>3%</td>
<td>4%</td>
<td>3%</td>
<td>2%</td>
</tr>
</tbody>
</table>

*Instructional context*

The students that participated in this study were enrolled in the same general chemistry lab course offered at neighboring institutions. Due to the common course numbering system used within the state where this study took place, both courses were aligned with common course standards. Furthermore, as identified in the syllabus of each course, both institutions identified similar course objectives. Table 2.2 compares the course objectives from each institution. Additionally, both courses followed a 15-week timeline, which included eleven investigations for both the comparison group and treatment group. Students in each course met for class one day each week.
Table 2.2 Course objectives, student outcomes for each group

<table>
<thead>
<tr>
<th>Comparison Group</th>
<th>Treatment Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Develop scientific inquiry skills (i.e. hypothesis generation, observation, description, data manipulation, and evaluation of results)</td>
<td>• “Think scientifically” (i.e. critical in observations, quantitative in measurements, analytical in drawing conclusions, assess reliability of results)</td>
</tr>
<tr>
<td>• Develop technical skills and experience using scientific instrumentation</td>
<td>• Introduce students to techniques and instrumentation commonly used in the scientific community</td>
</tr>
<tr>
<td>• Illustrate principles presented in lecture</td>
<td>• Reinforce concepts presented in lecture</td>
</tr>
<tr>
<td>• Promote problem-solving skills through collaborative and individual work</td>
<td></td>
</tr>
<tr>
<td>• Improve scientific writing skills through weekly reports</td>
<td></td>
</tr>
</tbody>
</table>

The two courses have similar structural and organizational features, however, there are marked differences between the instructional contexts of each course. The instructional model for the comparison group is identified as the traditional model, whereas the instructional model for the treatment group is argument-driven inquiry, the instructional model being investigated in this study. The argument-driven inquiry instructional model will be described in the following section; subsequently the two instructional approaches will be compared to show how they differ.

**Argument-driven inquiry (ADI)**

As discussed in the previous chapter, the argument-driven inquiry instructional model is designed to give students an opportunity to design an investigation, generate and evaluate data, develop arguments supporting a claim, critique the arguments of others, and generate a written report to articulate how their investigation and resulting argument adequately answers the research question that guided their investigation. The ADI approach is unique in the way the various phases of the process are structured to allow for a more authentic and educative experience for the students; the eight phases of ADI are presented and described in Table 2.3.
Table 2.3 *Phases of the ADI instructional model* (Sampson, Walker, and Grooms, 2009, p. 43)

<table>
<thead>
<tr>
<th>Phase</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification of a task</td>
<td>Teacher identifies task, phenomenon or problem to be solved</td>
</tr>
<tr>
<td>Generation and Analysis of data</td>
<td>Small groups of students generate data using a method of their own design</td>
</tr>
<tr>
<td>Production of a tentative argument</td>
<td>Each group articulates and justifies its explanation on whiteboards to be shared with other groups</td>
</tr>
<tr>
<td>Argumentation session</td>
<td>Each group shares their argument with the other groups for critique and then refines their argument</td>
</tr>
<tr>
<td>Investigation report</td>
<td>Individual students prepare a report explaining the goal of the work, the methods used, and a well-reasoned argument addressing the original task</td>
</tr>
<tr>
<td>Double-blind peer review</td>
<td>Investigation reports are reviewed by classmates according to rubrics to provide feedback to their peers</td>
</tr>
<tr>
<td>Revision of the report</td>
<td>Investigation reports are revised based on peer review process</td>
</tr>
<tr>
<td>Explicit and Reflective discussion</td>
<td>Related to the original task and content goals for the investigation</td>
</tr>
</tbody>
</table>

In order to complete the phases of the ADI instructional model each ADI investigation spans more than one class meeting for the treatment group. Typically one class meeting focuses on the data collection, data analysis and argumentation sessions. The following class meeting focuses on the peer review of the investigation reports. Once the peer review process is completed the students begin the investigation design phase for the next ADI activity or compete a non-ADI activity. These non-ADI activities are targeted at introducing students to a new piece of laboratory equipment, laboratory skill, or providing background experiences for the next ADI investigation.

*Comparison of instructional contexts*

Although several of the course objectives are similar for the comparison and treatment groups and the science content that is addressed in each course is similar, there are distinct differences between the type of instruction the students experience at each institution. Undoubtedly, each institution has selected the instructional methods they feel
are most appropriate in light of the various goals, objectives and outcomes they have for their particular students. It is also worth noting that neither institution has as a goal for their students, the ability to craft an argument supporting a stance related to a socioscientific issue. Therefore, it is not the intention of this study to make claims about the appropriateness of either instructional model in relation to the espoused goals of the courses, rather it is to simply identify if there are differences in the ways the students engage with socioscientific issues that may be related to the type of instruction experience by the students.

In order to clearly differentiate the two instructional approaches of the comparison and treatment groups, some common aspects of each course were identified through classroom observation. Table 2.4 indicates how those specific features manifest themselves within each instructional model. Key differences between the instructional models appear in the investigation design aspects and the nature of the student groups.

For each of the lab investigations within the comparison group, or traditional instructional model, the investigation procedures are provided to the students; whereas for the treatment group, or ADI model, the students must generate their own investigation procedures for the six ADI investigations. The treatment group students are provided procedures for the five non-ADI activities. Another difference between the instructional models is the number of students within each group. Within the treatment group, students

<table>
<thead>
<tr>
<th>Feature</th>
<th>Traditional (comparison group)</th>
<th>ADI (treatment group)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of labs</td>
<td>11 investigations</td>
<td>6 ADI investigations</td>
</tr>
<tr>
<td>Who generates investigation question/task?</td>
<td>Instructor</td>
<td>Instructor</td>
</tr>
<tr>
<td>Who develops investigation procedures?</td>
<td>Instructor</td>
<td>Students (ADI labs)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Instructor (non-ADI Labs)</td>
</tr>
<tr>
<td>Number of students per group</td>
<td>1 (8 investigations)</td>
<td>3 or 4 (all investigations)</td>
</tr>
<tr>
<td>Average time in lab (minutes)</td>
<td>95 of 180 scheduled</td>
<td>110 of 110 scheduled</td>
</tr>
</tbody>
</table>
work in teams of three or four (depending on class size) for all eleven investigations. Students in the comparison group work alone more often than not, however on some investigations the students do work with a partner.

Table 2.4 identifies similarities and differences within features common to both instructional approaches. There are additional differences that arise as well due to the specific features found only within the ADI model. For instance the comparison group does not engage in the peer review process, thereby the students do not receive feedback from one another regarding their reports. Additionally, the traditional instructional model does not provide an explicit opportunity for students to critique the data and claims of other students. The syllabus for the students in the traditional model does not suggest discussing data and conclusions with classmates, yet those opportunities are not specifically structured to ensure they take place. Based on the features described above for each instructional approach, it is evident that the instruction experienced by each group of students is indeed different.

**Procedures**

The students in this study were presented with two different SSI tasks, the EPA task and the Sugar task, which took the form of a short narrative followed by two competing television commercials that supported opposing views in regards to the action that should be taken to address the SSI. Both of the SSI tasks (described in more detail below) are situated in the larger context of proposing governmental action in an attempt to resolve some issue associated with the SSI (e.g. carbon emissions and health concerns). After the students read the narrative and watched the commercials for each task they completed a follow-up questionnaire. The students completed these two tasks at the outset of the semester and then once again at the close of the semester in order to determine if there were any changes from before to after the intervention. Within each group of students, approximately half of the students completed the EPA task first with the other half completing the Sugar task first. The tasks were arranged this way in an attempt to control for the possibility of students learning from any one particular task.

Due to the potential influence of students’ content knowledge, either initial knowledge or knowledge acquired during the intervention, the students completed a
multiple choice content knowledge assessment related to each SSI task pre- and post-intervention. This content assessment was completed prior to the students viewing the SSI tasks. Along with the content assessment the students completed an epistemological sophistication questionnaire pre- and post-intervention. The epistemology questionnaire allows for measurements of the students’ initial epistemological sophistication as well as any changes during the course of the 15-week intervention.

To facilitate the data collection process each SSI task and data collection instrument was housed on a website created specifically for this study. Using such a medium allowed the students to work at his or her own pace and read or view any of the available materials in whatever order they choose and as many times as were necessary for the individual student. The students accessed the website via computers provided by the researcher. Access to the Internet dictated how the computers were used, in that students within the treatment group completed the tasks in their original classroom, whereas students in the comparison group visited a computer lab within the same building (one floor above) as their original classroom. Each class section completed the tasks during their regularly scheduled class meeting times, either before or after other class activities as dictated by the individual instructors.

**SSI Tasks**

In this study the students were presented with two SSI tasks, one addressing the Environmental Protection Agency (EPA) placing carbon emissions limits on stationary sources (i.e. power plants) [referred to as the EPA task], the other addressing the levying of a tax on sugary beverages (i.e. carbonated sodas, energy drinks, etc.) [referred to as the Sugar task]. For the EPA task the students were provided with a description of actual legislation proposed by Senator J. Rockefeller (D-WV) that suggested the EPA should not place carbon emissions limits on stationary sources without providing congress ample time to address the potential economic impacts that such limits may have [see Appendix A for actual narrative]. In reference to the proposed legislation Sen. Rockefeller stated,

> This legislation will issue a two-year suspension on EPA regulation of greenhouse gases from stationary sources—giving Congress the time it needs to address an issue as complicated and expansive as our energy future. Congress, not the EPA, must be the ideal decision-maker on such a challenging issue (Rockefeller, 2010).
Sadler (2004) describes SSI as ill-structured problems that require the coordination of a variety of perspectives in order to come to a decision on how to act regarding the specific issue. Limiting carbon emissions from stationary sources, as described in the EPA task, fits this description of an SSI. Enacting such limits may have positive effects for the environment, while at the same time have negative impacts on companies or perhaps an industry as a whole. Additionally, individuals working within these industries may lose their jobs as a result of such limits; conversely, new jobs may also be created to address technological needs to conform to the new regulations. Finally, this proposed legislation also highlights the responsibilities of various stakeholders within the government. Senator Rockefeller is charged to act on behalf of his constituents and the EPA is charged with protecting the environment; both of these stakeholders must try to accomplish their agendas while also accommodating other viewpoints in order to be successful.

Following the narrative that described the legislation, students were presented with two TV commercials addressing energy use; each commercial was thirty seconds long. One commercial was paid for by Repower America™; an organization that supports the use of alternative energy sources and the reduction of dependence on fossil fuels to supply the energy needs of the U.S., because of environmental concerns and for economic growth. The second commercial was paid for by the American Coalition for Clean Coal Electricity™; an organization that supports the continued use of coal as an energy source because of its efficiency and economic feasibility. Both commercials were readily available on the Internet and downloaded to the project website from YouTube™.

The Sugar task provided students with an article written by the Center for Science in the Public Interest (CSPI) that supported the levying of a tax on sugary beverages like non-diet sodas, energy drinks, sports drinks, etc [see Appendix B for the actual narrative]. In the article the CSPI suggests there is a link between increased consumption of high-calorie beverages and obesity and further suggests “soft drink taxes provide a dual benefit; they could drive down soft drink consumption and generate

---

1 Repower America commercial downloaded from: http://www.youtube.com/watch?feature=player_embedded&v=GlQc9Kj15NM
2 American Coalition for Clean Coal Electricity commercial downloaded from: http://www.youtube.com/watch?feature=player_embedded&v=SHxykaAoTvc
revenue for much-needed health programs” (CSPI, 2010, para. 2). A key component of
the suggestions by the CSPI is that health problems related to obesity create a burden on
health systems (i.e. Medicare and Medicaid) and ultimately are paid for by tax payers at
large, whereas a “soda tax” could help to offset such burdens by increasing revenue and
reducing consumption. Much like the EPA task, this task requires students to consider
multiple viewpoints as to the justification of such a tax; including but not limited to the
beverage companies that may see reduced profits, health organizations that may see
increased funding, individuals whose health may improve, etc.

As in the EPA task, following the short narrative for the Sugar task, the students
were presented with two TV commercials that offered differing views on the enactment
of such a tax; again, each commercial was thirty seconds long. The specific commercials
used in this task were in reference to levying a “sugar tax” in New York State. One
commercial³ is paid for by the American Beverage Association, which opposes the tax
and uses a local small-business owner to highlight the potential negative effect on his
customers. The other commercial⁴ is paid for by Alliance for a Healthier New York,
which is in favor of the tax and uses health professionals to highlight the potential health
benefits of such a tax. Both of these commercials were also available via YouTube™ and
downloaded directly to the project website.

After reading the text and watching the commercials associated with each SSI
task the students completed a follow-up questionnaire. The students were asked to
indicate how they would vote (“in favor of”, or “in opposition to”) with respect to the
legislation proposed in each task. Following their vote the students were asked to provide
an argument supporting why they voted the way they did. Finally, the students were
asked to identify the extent to which a variety of factors contributed to their vote, such as
their morals and values, economic considerations, what they know about science,
political views, etc.

³ American Beverage Association commercial downloaded from:
http://www.youtube.com/watch?feature=player_embedded&v=Uz5TOMUhnP0
³ Alliance for a Healthier New York commercial downloaded from:
http://www.youtube.com/watch?feature=player_embedded&v=2sLANBAUNp0
These two socioscientific issues were chosen because of their potential to engage the student participants. Defining these two issues as SSI implies that there are many complex facets to developing a resolution to the particular issue and many viewpoints must be taken into account, such as science, politics, and economics. Both the EPA task and the Sugar task incorporate these factors and are relevant to the students’ personal lives. By using the TV commercials the students are presented with a media type they are likely very familiar with and encounter on a daily basis. The very nature of advertisements are to convince the viewer of a particular stance or claim and in the context of this study the TV commercials represent opposing views to be considered by the participants.

Data Collection Instruments

Given the discussion above and the research questions guiding this study, it is important to assess the students’ content understanding of the science topics that underlie the SSI they will be asked to consider, the students’ epistemological sophistication both pre- and post-intervention as well as the quality of their arguments supporting their stance on the two SSI tasks. Therefore, all students completed an initial survey instrument that assessed their understanding of scientific topics related to the issues addressed in the EPA task and the Sugar task (i.e. fossil fuels, energy, nutrition, and health), and their initial level of epistemological sophistication (devised from; Kuhn, Cheney & Weinstock, 2000). The students completed each data collection instrument both pre- and post-intervention.

Science content knowledge assessment

The content knowledge assessment is a series of multiple choice items gathered from a variety of standardized instruments. The total assessment consisted of twenty-one items, ten related to the Sugar task and eleven related to the EPA task. Reliability of the assessment was conducted using pilot test data completed during the summer semester prior to this study. The results of that analysis revealed a KR-20 score of .66 for the items related to the Sugar task and a KR-20 score of .65 for the items related to the EPA task. According to Salvucci et al. (1997), reliability statistics within the range of 0.5 to 0.8 are considered moderate. Furthermore, when the purpose of an assessment is to differentiate
groups, as is the case in this research, reliability measures as low as 0.6 are deemed acceptable (Fitz-Gibbon and Morris, 1987). Since the content knowledge assessment is only being used to differentiate the content knowledge of the comparison and treatment groups (not individual students), the reliability statistics above are appropriate, indicating sufficient reliability for the content knowledge assessment in the context of this study [see Appendix C for the content knowledge assessment].

**Epistemological sophistication questionnaire**

An epistemological sophistication questionnaire adapted from Kuhn, Cheney, and Weinstock (2000) was used to assess the students’ epistemological sophistication. The original assessment addressed students’ ideas about knowledge on five sub-scales: judgments of personal taste, aesthetic judgments, value judgments, judgments of social truth, and judgments of physical truth (Kuhn et al., 2000). Within each knowledge domain there are three pairs of statements the students read and then decide whether one statement could be correct, both could be correct, or neither is more correct than the other. For the purposes of this study the most relevant sub-scales are those addressing value judgments, judgments of social truth, and judgments of physical truth; therefore, those were the only items used in the assessment of students’ epistemological sophistication. Below is an example of a pair of statements and possible responses from the value judgments sub-scale [see Appendix D for the epistemological sophistication assessment].

Robin thinks people should take responsibility for themselves.
Chris thinks people should work together to take care of each other.

A. Only one is right.
B. Both could have some rightness, but one could be more right.
C. Both could have some rightness, but one could not be more right than the other.

Unlike the science content assessment there are no right or wrong answers for the items contained within the epistemological survey. The students had to read several pairs of sentences and then evaluate whether one statement should be considered correct, both statements have some correctness with neither more correct than the other, or both statements have some correctness with one being more correct than the other. Each response provides insight to the students’ epistemological sophistication with respect to that particular set of statements.
This assessment consists of three sub-scales (judgments about values, social truth, and physical truth) each containing three sets of sentences (i.e. three questions per sub-scale) for the student to evaluate. In the original assessment by Kuhn, Cheney, and Weinstock (2000) the answer choices were slightly different, consisting of an initial question about whether or not one statement was right or if both could be right, with a follow-up question contingent on the initial answer. Because this assessment was completed using web-based survey software, duplicating the original format was not the most feasible approach, due to software issues, and would have been cumbersome for the student participants. Nussbaum, Sinatra, and Poliquin (2008) encountered similar formatting issues when they used the same assessment with web-based software. The answer format above reflects their adaptation to the answer choices of the original assessment, which addresses the formatting issue associated with the web-based approach; their format was also used in this study.

The restructuring of Kuhn, et al.’s (2000) original instrument may raise concerns about the distribution of students among the three epistemological categories of absolutist, multiplist, and evaluativist. Specifically, a concern may arise that students would be able to discern the “right” answer given that all answer choices were presented at once (as opposed to Kuhn, et al.’s binary choice format with a follow-up question) and the total number of choices was reduced from four to three. If this concern manifests itself, the result would be an increase in the number of students categorized as evaluativist.

Based on an analysis of pilot study data using the restructured format for the epistemology instrument, I would contend that the restructuring of the instrument does not result in an increase of the number of students categorized as evaluativist when compared to the original instrument. Of the forty-four students that participated in the pilot study only ten students (23%) were categorized as evaluativist when taking into account all three sub-scales assessed (value judgments, social truth judgments, and physical truth judgments). Furthermore, only four of those students (9% of the total sample) chose the evaluativist answer choice on each item.

A meta-analysis of data presented from the initial validation of the epistemology instrument (Kuhn et al., 2000) is shown below in Table 2.5. This table shows the
distribution of participants by epistemological category across the three judgment domains in the pilot study. The data below show the similarities in distribution for the two studies even though the instrument had a slightly different structure. It is important to remember, however, that the analysis below includes all sub-groups that participated in the original validation efforts by Kuhn, et al. (2000) ranging from fifth graders, through undergraduates, to professionals and even includes experts. My pilot study, on the other hand, only targeted community college students. Therefore, a second set of comparisons is needed.

Table 2.5 Distribution of participants by epistemological category across three judgment domains in Kuhn, et al. (2000) and Grooms pilot studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Absolutist</th>
<th>Multiplist</th>
<th>Evaluativist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kuhn, et al. (all sub-groups, n=107)</td>
<td>5</td>
<td>67</td>
<td>32</td>
</tr>
<tr>
<td>Grooms, pilot (community college, n=44)</td>
<td>2</td>
<td>75</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 2.6 below, shows the percentage of participants within similar sub-groups that were categorized as evaluativist on each of the three judgment domains. In the original study there were two sub-groups included that may be similar to the group targeted by my pilot study. Kuhn, et al. (2000) included undergraduate students from a highly selective private university which were included in the study because they represented “high intellectual ability but limited life experience” (p. 315). A second sub-group they included was community college students mostly enrolled in vocational programs that were assumed to be of lesser intellectual ability than the undergraduates. Based on these vague descriptions in the original study, I believe it is safe to assume that the students that participated in my pilot study would represent a group of students somewhere between these two extreme sub-groups used by Kuhn, et al. (2000). Therefore, Table 2.6 shows each sub-group as well as an average of the two extreme sub-groups used in the original study as a basis for comparison to my pilot study data.
When the data from my pilot study are compared to the average of Kuhn, et al.’s (2000) similar sub-groups there are comparable distributions of students categorized as evaluativist. Most comparable are the results for the value judgment domain, followed by the physical truth domain. There is some difference found in the social truth domain, however, this difference is difficult to tease apart given the likely variation between the social backgrounds of these sub-groups. Viewed holistically, the data presented here support the use of the restructured epistemology instrument. The concern that there would be a dramatic shift in the number of evaluativist based on students’ abilities to pick out the “right” answer was not borne out by these data.

Table 2.6 Percentage of participants categorized as evaluativist on three different judgment domains (Kuhn et al., 2000)

<table>
<thead>
<tr>
<th>Study</th>
<th>Percent Evaluativist</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Values Judgments</td>
</tr>
<tr>
<td>Kuhn, et al. (undergraduate sub-group, n=20)</td>
<td>45</td>
</tr>
<tr>
<td>Kuhn, et al. (community college sub-group, n=20)</td>
<td>25</td>
</tr>
<tr>
<td>Kuhn, et al. (average of undergraduate and community college sub-groups, n=40)</td>
<td>35</td>
</tr>
<tr>
<td>Grooms, pilot (community college, n=44)</td>
<td>36</td>
</tr>
</tbody>
</table>

**SSI follow-up questionnaire**

Finally, the quality of the students’ arguments supporting their stance on the SSI tasks was assessed. After the students read the narrative associated with each SSI and watched the accompanying TV commercials they completed a follow-up questionnaire. The questionnaire first asked them to identify how they would vote with respect to the SSI task they just experienced. For the EPA task, the prompt was:

If you, as a citizen, were given the opportunity to vote for or against the legislation proposed by Senator Rockefeller, how would you vote?

In FAVOR of the legislation (YES) [meaning you support the two year ban on the EPA regulations]
In OPPOSITION to the legislation (NO) [meaning you do not support the two year ban on the EPA regulations]

After identifying their stance on the issue, the students were asked to provide an argument that supported the vote they just cast. Once the students provided an argument supporting their vote on the SSI they were finally asked to identify to what extent certain factors influenced their vote. The various factors included: their political views, morals and values, economic considerations, what they know about science in general, and what they know about science content specific to the individual SSI tasks (i.e. fossil fuels, carbon emissions, or health and nutrition).

The follow-up questionnaires to each SSI also contained several other questions to elicit students’ ideas about data and evidence, what pieces of evidence from the tasks were influential, what additional information would they want as part of this task, etc. These questions do not directly address the research questions guiding this study and were not analyzed as part of this project; however, they may be useful for secondary analyses in regards to additional important research questions that are outside the scope of this initial exploratory investigation [see Appendices E and F for the SSI follow-up questionnaires].

**Data Analysis**

The following sections describe the scoring processes and rubrics used for each of the data collection instruments identified above. Actual data and results from this study, including sample arguments and justifications, will be presented in chapter three.

**Science content knowledge**

The science content knowledge assessment consisted of multiple-choice questions, this type of assessment allowed for the quantitative comparison of the treatment and comparison groups pre- & post-intervention. Independent samples t-tests were used to determine whether or not the two student groups have significantly different content knowledge related to the science topics underlying the SSI tasks that were used for collecting the argument data. Given the electronic nature of the data collection process, the assessment items were automatically scored in a dichotomous fashion as either right or wrong. In order to make comparisons between students with high and low
content knowledge a median split approach was taken. As such, the students scoring consistent with the median score and below were categorized as having low content knowledge and those scoring above the median level were categorized as having high content knowledge.

**Epistemological sophistication**

The assessment of epistemological sophistication provided a means to determine if students’ epistemological sophistication changed from the pre-intervention to post-intervention assessment. This assessment also helped to identify any initial similarities or differences between the treatment and comparison groups. Based on the responses provided by the student, he or she was identified as having an absolutist, multiplist, or evaluativist stance on each question; combining the three responses within each sub-scale provided a level of epistemological sophistication for that knowledge domain. Table 2.7 below shows how the various responses and combinations within sub-scales were coded and subsequent epistemological sophistication determined.

In order to determine the overall level of epistemological sophistication, each response provided by the student was scored according to the values presented in Table 2.7. Next, a total score was determined for each sub-scale based on the sum of the scores for each item within that subscale. Sub-scale scores of 0 or 1 correspond to an absolutist level with respect to that sub-scale, scores from 2 through 4 correspond to a multiplist level, and a score of 5 or 6 correspond to an evaluativist level. Once an epistemological category was identified for each sub-scale an overall level of epistemological sophistication was determined. Each epistemological category, at the sub-scale level, was scored 0 if absolutists, 1 if multiplists, and 2 if evaluativist. Finally the sub-scale scores were totaled and the resulting sum indicated the overall level of epistemological sophistication. Again, the score range for each overall epistemological category corresponded to those described above and identified in Table 2.7. Additionally, Table 2.7 identifies sample combinations of sub-scale epistemology categories that would result in specific overall epistemological sophistication categories. By using non-parametric Chi square analyses, the distributions of students among the various epistemological levels can be compared within and between groups as well as pre- and post-intervention.
Table 2.7 *Coding scheme for epistemological sophistication within sub-scale*

<table>
<thead>
<tr>
<th>Level of epistemological sophistication (ES)</th>
<th>Corresponding response</th>
<th>Corresponding code value</th>
<th>Possible total sub-scale score for level of ES</th>
<th>Possible response combinations within sub-scale or across sub-scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolutist (A)</td>
<td>Only one is right.</td>
<td>0</td>
<td>0, 1</td>
<td>AAA, AAM</td>
</tr>
<tr>
<td>Multiplist (M)</td>
<td>Both could have some rightness, but one could not be more right than the other.</td>
<td>1</td>
<td>2, 3, 4</td>
<td>AMM, AAE, MMM, AME, MME, AEE</td>
</tr>
<tr>
<td>Evaluativist (E)</td>
<td>Both could have some rightness, but one could be more right.</td>
<td>2</td>
<td>5, 6</td>
<td>MEE, EEE</td>
</tr>
</tbody>
</table>

*Argument quality*

*Argument structure*

Informal reasoning is suggested as the process underlying the construction and evaluation of arguments, especially when pertaining to ill-structured problems such as SSI (Means and Voss, 1996; Sadler, 2004a). Student progress toward scientific literacy, as discussed in the previous chapter, is related to the improvement in their abilities to use informal reasoning in the context of socioscientific issues. One inferential means of evaluating a student’s informal reasoning abilities is the evaluation of the structure of the arguments he or she produces. Evidence indicating an increase in the quality of students’ arguments might suggest that the instructional strategies investigated in this study lead to enhanced scientific literacy.
The evaluation criteria for the arguments students produced in this study draw on the work of Means and Voss (1996), Sadler and Donnelly (2006a), Zohar and Nemet (2002), and Venville and Dawson (2010). These frameworks focus on key aspects of informal reasoning and emphasize the inclusion of multiple reasons supporting the argument, a rationale as to why those reasons are included, consideration of multiple perspectives, the acknowledgment of counter-arguments, and the quality of the supporting reasons used. Collectively these criteria address the argument’s structure and the nature of the justifications the students use as support for their arguments.

Table 2.8 shows the rubric used in this study to evaluate the structure of students’ arguments based on these criteria. In the SSI tasks, the students’ vote served as the claim for their argument, therefore no criteria related to a claim were included in this rubric. Along with that notion, is the potential score of zero for the students’ argument structure, if a students’ supporting argument does not agree with their actual vote, then there is no internal consistency for the argument, thereby the argument received a score of zero. The inclusion of more sophisticated argument structures, such as rationales and rebuttals increases the score of the argument as these types of structures increase the overall sophistication of the argument.

<table>
<thead>
<tr>
<th>Argument Level</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Argument does not support vote, or there are no reasons provided to support vote</td>
</tr>
<tr>
<td>1</td>
<td>Single or multiple reasons provided to support vote</td>
</tr>
<tr>
<td>2</td>
<td>Single or multiple reasons provided to support vote, WITH a rationale</td>
</tr>
<tr>
<td>3</td>
<td>Single or multiple reasons provided to support vote, WITH a rebuttal</td>
</tr>
<tr>
<td>4</td>
<td>Single or multiple reasons provided to support vote, WITH a rationale, WITH a rebuttal</td>
</tr>
</tbody>
</table>

The student participants in this study provided a supporting argument for two different SSI tasks, both pre- and post-intervention, resulting in approximately 600 total arguments. The researcher and a second rater used the rubric in Table 2.8 to individually
score a random sub-sample of approximately 10% of all the arguments, which resulted in an 80% agreement in argument structure between the two raters. Additional inter-rater reliability assessments indicated a Cohen’s Kappa value of 0.71, indicating good agreement using this argument structure rubric. The raters negotiated any disagreements in score and a consensus score was given; both raters scored all of the remaining arguments and consensus scores were given in each case. Furthermore, the scoring of each argument was conducted in such a manner that the two raters were blind to the condition (i.e. comparison or treatment group) and the time (i.e. pre- or post-intervention) for each argument.

**Nature of the argument justifications**

The second aspect of the argument that was evaluated was the nature of the justifications. This aspect of the argument was included as a means of assessing what types of information the students seemed to privilege in regard to the decision the students made or argument they constructed related to the SSI. For example the nature of a particular justification might be scientifically oriented, politically motivated, or grounded in economics; all of these factors or considerations, and others, could play a role in the decisions made in order to resolve the complex SSI used in this study. It is hoped that students developing scientific literacy during a given course would more readily use their understanding of scientific practice and habits of mind to inform the decisions (and arguments) they make relative to the SSI. Such a shift in thinking would become apparent in the increased use of scientific arguments supporting their stance on the SSI discussed in this study. Given that neither curriculum investigated in this study specifically addresses the underlying science content of the SSI tasks it is unlikely that increases in the scientific nature of the students’ justifications will be evident by increases in their use of science content knowledge. It is more likely that other aspects of scientific thinking will prevail, such as support or critique based on methodological grounds and appeals to research or evidence when developing arguments. Table 2.9 below identifies the types of justifications coded for in this study.

The justifications identified in Table 2.9 were generated from the literature related to students’ engagement with SSI tasks (Sadler et al., 2007; Venville and Dawson, 2010) as representing typical types of justifications used by students. The researcher and a
second rater used the rubric in Table 2.9 to individually categorize a random sub-sample of approximately 10% of all the arguments based on justification type, which resulted in 77% agreement in classification between the two raters. Additional inter-rater reliability assessments indicated a Cohen’s Kappa value of 0.71, indicating good agreement using the justification rubric. The raters negotiated any disagreements in score and a consensus score was given; both raters categorized all of the remaining justifications and consensus scores were given in each case. Furthermore, the identification of justifications was conducted in such a manner that the two raters were blind to the condition (i.e. comparison or treatment group) and the time (i.e. pre- or post-intervention) for each argument.

Table 2.9 Types of argument justifications and descriptions

<table>
<thead>
<tr>
<th>Justification</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intuitive</td>
<td>Based on the notion of “that makes the most sense”, “everyone agrees”, etc.</td>
</tr>
<tr>
<td>Emotive</td>
<td>Expresses care, sympathy, empathy, concern for those affected, etc.</td>
</tr>
<tr>
<td>Greater good</td>
<td>Focuses on a benefit to society, overall health, cleaner environment, etc.</td>
</tr>
<tr>
<td>Economic</td>
<td>Focuses on cost, impact on the number of jobs, generate revenue, etc.</td>
</tr>
<tr>
<td>Scientific explanation or Empirical</td>
<td>References a scientific theory, model, law, principle, or concept, or References findings from a study, methodological issues, etc.</td>
</tr>
<tr>
<td>Political</td>
<td>Focuses on the role of government</td>
</tr>
</tbody>
</table>
CHAPTER THREE

Results and Discussion

Influence of various factors on students’ votes related to SSI

As discussed in a previous chapter, socioscientific issues are complex problems that often involve scientific, social, political, and economic aspects. When an individual is tasked with making a decision related to an SSI, such as casting a vote on a particular piece of legislation, he or she must weigh the many aspects of the issue and then make a decision based on the factors that are deemed most relevant or important. The available literature suggests that the above factors play a role in this decision-making process (Sadler et al., 2007), but are some of these factors privileged more than others by undergraduates?

The participants in this study were asked to rate how various factors, such as politics, economics, science, and morals, influenced their vote on each of the SSI tasks they were presented. Each factor was rated on a scale from zero to four, with zero representing no influence on their vote and four representing “a great deal” of influence. For the initial analysis of the responses, the two tasks and two groups were combined to provide an overview of how undergraduate students rated these factors. Table 3.1 shows the results of this initial analysis. Before discussing the student rankings of these factors it is important to note that this data source is based on student self-report regarding the influence of each factor. The self-report nature of this data does not render it worthless, however, one must be cautious of such data and mindful of how the data may have been influenced. Particularly, consider that the data were collected within a science class, which may have had the unintended consequence of students inflating the claimed influence of scientific ideas and similarly deflating the claimed influence of political ideas under the guise of “pleasing the researcher” or providing what they thought was the “right” answer. It is with these limitations in mind that the data related to this aspect of the study are presented.

These data show that the students in this study claimed their content specific science knowledge had the most influence on their vote with an average rank of 2.83. The content specific science prompt was different for each task; students were asked how
much their knowledge about climate change influenced their vote for the EPA task and how much their knowledge about health and nutrition influenced their vote for the Sugar task. Furthermore, the claimed high influence of this specific science content knowledge was significantly higher than any other factor. The factor claimed to have the least influence on their vote was ‘political views’ with a mean ranking of 1.20. This factor was ranked significantly lower than any other factor. Morals and values, economic considerations, and the general science knowledge factors were considered to have similar influence on the students’ votes. Although the students’ mean rankings differ slightly for these three factors, none of them were ranked significantly different from one another.

Table 3.1 *Mean rating for factors influencing students’ votes, pre-intervention*

<table>
<thead>
<tr>
<th>Factor</th>
<th>Mean Rating (SD)</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Morals &amp;</td>
<td>Economics</td>
</tr>
<tr>
<td>Content Specific Science</td>
<td>2.83 (1.22)</td>
<td>.21&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Morals &amp; Values</td>
<td>2.62 (1.35)</td>
<td>-</td>
</tr>
<tr>
<td>Economics</td>
<td>2.61 (1.30)</td>
<td>-</td>
</tr>
<tr>
<td>General Science</td>
<td>2.45 (1.26)</td>
<td>-</td>
</tr>
<tr>
<td>Politics</td>
<td>1.20 (1.41)</td>
<td>-</td>
</tr>
</tbody>
</table>

*significant at the .05 level

<sup>a</sup> t(303) = 2.36, p = .02  
<sup>b</sup> t(303) = 2.35, p = .02  
<sup>c</sup> t(303) = 5.52, p < .01  
<sup>d</sup> t(303) = 15.27, p < .01  
<sup>e</sup> t(303) = .14, p = .89  
<sup>f</sup> t(303) = 1.96, p = .05  
<sup>g</sup> t(303) = 16.30, p < .01  
<sup>h</sup> t(303) = 1.66, p = .10  
<sup>i</sup> t(303) = 14.08, p < .01  
<sup>j</sup> t(303) = 12.46, p < .01

In order to provide a more nuanced picture of how the students claim to have been influenced when making their votes, consider the more fine-grained analysis of the rankings between the comparison and treatment groups and between the two SSI tasks. Table 3.2 shows that both the treatment and comparison groups rated the influence of the various factors the same; not only do the two groups provide the same relative ratings for each factor, they also scored the factors with similar claimed strength of influence.
Table 3.2 Mean rating for factors influencing students’ votes, by group, pre-intervention

<table>
<thead>
<tr>
<th>Group</th>
<th>Content Specific Science</th>
<th>Morals &amp; Values</th>
<th>Economics</th>
<th>General Science</th>
<th>Politics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison</td>
<td>2.82 (1.25)</td>
<td>2.61 (1.36)</td>
<td>2.60 (1.33)</td>
<td>2.46 (1.28)</td>
<td>1.20 (1.40)</td>
</tr>
<tr>
<td>Treatment</td>
<td>2.84 (1.19)</td>
<td>2.62 (1.34)</td>
<td>2.61 (1.27)</td>
<td>2.45 (1.23)</td>
<td>1.21 (1.42)</td>
</tr>
</tbody>
</table>

*Note. 0 = No influence, 4 = Great influence*

However, when evaluating the influence of the various factors in the context of each SSI task, the students appeared to privilege different factors when making their vote (see Table 3.3). Although the students regarded politics and general science as the least influential factors in both contexts, economic considerations and content specific science knowledge seem to play different roles. For the EPA task, the students rated economic factors as the most influential while content specific science knowledge was ranked as the most influential for the Sugar task. Even though economics has the most influence on student votes for the EPA task, its influence is not significantly greater than the influence of morals and values, content specific science, or general science, nor is the economic influence in the EPA task significantly different from its influence on the Sugar task. On the other hand, content specific science knowledge, i.e. ‘what the student knows about health and nutrition,’ was identified as having the most influence on their vote. The mean rating of 3.07 for the Sugar task was significantly higher than any other factor for that task; furthermore content specific science knowledge had a significantly greater influence on students’ votes related to the Sugar task than it did for the EPA task.

An initial analysis of this type merely provides an overview of how the undergraduate participants in this study claimed to have been influenced when making their votes on the two SSI tasks they were presented. Based on these results one might expect the justifications used in the arguments supporting the students’ votes would echo the claimed influence of these factors, thereby resulting in supporting arguments grounded in content specific science, economics, and moral and value considerations, due to their claimed high influence. The following section describes the distribution of actual
justifications observed in the students’ arguments supporting their votes on the SSI-related tasks.

Table 3.3 Mean rating for factors influencing students’ votes, by task, pre-intervention

<table>
<thead>
<tr>
<th>Task and Factor</th>
<th>Mean Rating (SD)</th>
<th>Morals &amp; Values</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Content Specific Science</td>
</tr>
<tr>
<td>EPA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economic</td>
<td>2.64 (1.29)</td>
<td>.02^a</td>
<td>.05^b</td>
</tr>
<tr>
<td>Morals &amp; Values</td>
<td>2.62 (1.29)</td>
<td>-</td>
<td>.03^e</td>
</tr>
<tr>
<td>Content Specific Science</td>
<td>2.59 (1.26)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>General Science</td>
<td>2.55 (1.24)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Politics</td>
<td>1.25 (1.41)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sugar</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content Specific Science</td>
<td>3.07 (1.13)</td>
<td>.45^k*</td>
<td>-</td>
</tr>
<tr>
<td>Morals &amp; Values</td>
<td>2.62 (1.41)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Economic</td>
<td>2.57 (1.31)</td>
<td>.05^l</td>
<td>.50^m*</td>
</tr>
<tr>
<td>General Science</td>
<td>2.36 (1.26)</td>
<td>.26^p**</td>
<td>.71^q*</td>
</tr>
<tr>
<td>Politics</td>
<td>1.16 (1.41)</td>
<td>1.46^r*</td>
<td>1.91^s*</td>
</tr>
</tbody>
</table>

* significant at $p = .01$ level  
**significant at $p = .05$ level

a. $t(151) = .196, p = .85$  
b. $t(151) = .436, p = .66$  
c. $t(151) = .782, p = .44$  
d. $t(151) = 10.67, p < .01$  
e. $t(151) = .277, p = .78$  
f. $t(151) = .665, p = .51$  
g. $t(151) = 11.75, p < .01$  
h. $t(151) = .432, p = .67$  
i. $t(151) = 9.05, p < .01$  
j. $t(151) = 10.02, p < .01$

Nature of justifications supporting students’ votes on SSI

The participants provided a written argument supporting the vote they cast for each SSI task and those arguments were subsequently scored using the nature of justifications rubric presented in Table 2.9. By scoring the nature of the justifications the students used it was possible to tell whether or not the factors the students claimed influenced their votes were actually present in their supporting arguments. It was possible for the supporting arguments to contain more that one type of justification; each justification used by a student was scored and included in the analysis.
Overall, the 152 students provided 518 justifications for the two SSI tasks during the pre-intervention assessment. Table 3.4 shows how those justifications were divided among the comparison and treatment groups for each of the two tasks. Each group provided a similar number of justifications on each task, with the Sugar task generating slightly more justifications than the EPA task. Additionally, each group maintained a similar average number of justifications for each task. The average number of justifications statistic has both positive and negative potential interpretations. Providing more justifications can imply a well-constructed argument, which considers multiple factors, or alternatively, a less sophisticated argument that simply touches on a variety of disconnected reasons. Additional analysis is needed to differentiate this effect, which will be discussed below.

An analysis of the type of justifications the students used in their arguments indicates that 35% of the justifications used to support their votes were identified as promoting a sense of ‘greater good.’ Students making arguments of this nature supported their vote by expressing concern for society as a whole, general concern for the environment, or general concern for people’s health. This type of justification was the most common during the pre-intervention assessment. Figure 3.1 shows the overall distribution of justifications for the pre-intervention assessment of both SSI tasks for both the comparison and treatment groups combined. The second most common type of justification supporting students’ votes were those grounded in economic concerns, representing 24% of all justifications. The least common type of justification supporting students’ votes was scientific or empirically based reasons (examples of these justifications are included below in the context of each task).

Previously the students claimed content specific science knowledge, morals and values, and economic considerations had the most influence on their votes. The

<table>
<thead>
<tr>
<th>Task</th>
<th>Number of Justifications</th>
<th>Average Justifications per Student</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Comparison (N=79)</td>
<td>Treatment (N=73)</td>
</tr>
<tr>
<td>EPA (% by task)</td>
<td>117 (47%)</td>
<td>130 (53%)</td>
</tr>
<tr>
<td>Sugar (% by task)</td>
<td>135 (50%)</td>
<td>136 (50%)</td>
</tr>
</tbody>
</table>
distribution of justifications displayed in Figure 3.1, however, does not reflect how these students rated these various factors. Overall, the students touted content specific science knowledge as the most influential factor affecting their vote, yet only 4% of the total justifications used by the students are scientific or empirical in nature. Political views were claimed to have the least influence on students’ votes, yet 12% of all the justifications were political in nature. Despite students claiming that content specific science had over twice the amount of influence (mean rank of 2.83) on their vote when compared to political views (mean rank of 1.20), there were three times as many political justifications than scientific.

![Distribution of Justifications, Pre-Intervention](image)

*Figure 3.1 Overall distribution of pre-intervention justifications for both tasks and both groups*

The students also ranked morals and values and economic considerations as having a high influence on their vote. This claim is further supported by the prevalence of observed morals and values-based and economic-based justifications used by the students when supporting their vote. Justifications consistent with morals and values (i.e. justifications grounded in a notion of ‘greater good’) represent 35% of the total
justifications and economic justifications represent 24% of all justifications pre-intervention.

SSI specific justifications

As discussed above, there were slight differences in the influence of various factors on students’ votes based on the context of the SSI. Therefore, it is essential to disaggregate the justifications used by the students to examine how the context of their votes and supporting justifications may be linked. Figure 3.2 displays the distribution of justifications by SSI task. The types of justifications have the same general distribution when separated by task, with the one exception of political and intuitive justifications within the Sugar task. Intuitive justifications represent 19% of the justifications for the Sugar task and only 11% for the EPA task, making intuitive justifications the third most common type for the Sugar task and fourth most common type for the EPA task. A possible explanation for this is the personal nature of the Sugar task. In that task the students were deciding whether or not they were in favor of a tax on sugary soft drinks which could have a more direct influence on their personal lives than the EPA task which represents an issue more removed from the individual students’ lives.

In the case of the Sugar task the intuitive arguments focused on personal choice and how the individual would be affected by the implementation of the tax. One student provided the argument:

Consuming soft drinks is a choice and a right. I consume soft drinks just not in excess. I'm not overweight so therefore I do not fit the reasoning for taxing soft drinks. (Student, T001)

While another stated:

I shouldn’t be penalized by paying more for a soft drink or juice drink just because other people are obese, since not all of us are (Student, T003)

These students are focused on the personal impact the tax would have on them as opposed to the overarching strengths or weaknesses of the proposed tax. Alternatively, the EPA task evoked intuitive arguments of a different nature. These intuitive arguments

---

5 Student ID numbers denote group membership and pre- or post-condition. “T” represents treatment group members and “C” represents comparison group members. The first number in the three-digit code indicates pre- or post-condition, with 0 representing pre- and 1 representing post. The final two digits are unique to individual students.
follow a rhetorical notion along the lines of a “sounds good to me” attitude. Consider the following example arguments:

I believe the ban is a good idea as it will help to formulate new ideas in that time in order to try something different. (Student, C002)

and,

If we can do something to protect the planet that we live on, why shouldn't we? (Student, T002)

Neither of these students expresses a strong personal concern regarding the proposed legislation involved in this task as was seen in the intuitive arguments for the Sugar task.

Figure 3.2 Distribution of justifications for each SSI task, pre-intervention

For both tasks, justifications grounded in a sense of ‘greater good’ represent the largest percentage of justifications for that task. In the context of the EPA task 41% of the justifications were categorized as expressing concern for the greater good, usually in this task with regard to the environment in general. Consider the following examples:

I would vote against the two year ban because it would cause pollution. (Student, C003)
Justifications supporting the notion of a ‘greater good’ were common in the supporting arguments for the Sugar task as well, accounting for 31% of all the justifications. In the context of the Sugar task, justifications of this nature typically promoted the enhanced health of individuals or a general concern for healthy living. The examples below demonstrate this trend.

Majority of Americans are over weight and if taxing the drinks are the only way to help their health then I think it is a good idea. (Student, C005)

Though I would not like to pay more for a soda because of the tax, it will help a great deal in slowing the tendency of obesity. (Student, T005)

The second most common type of justification focused on economic concerns. This type of justification represented 23% and 25% of the justifications for the EPA and Sugar tasks respectively. Due to the differences in the contexts of the two tasks the focus of the economic justifications differs between the two. A typical economic justification for the sugar task reads as follows:

The tax could lower the obesity of millions of Americans and could also provide a way for the government to increase national revenue. (Student, C008)

The Sugar task is, in part, an economic issue in that it is proposing a new tax; the EPA task relates to the economy based on a concern for job retention. Therefore, the economic justifications tend to focus on generating revenue for the Sugar task or protecting jobs in regard to the EPA task. The following argument illustrates how economic concerns influenced this student’s vote on the EPA task:

I would vote in favor of a two year ban on EPA regulations only because in these desperate economic times, jobs and the rising unemployment rates throughout the country need to take priority over what the EPA may deem dangerous to the environment. Though a safe environment is highly important, it is also extremely essential for the government to protect jobs at this moment. (Student, T007)

When the participants in this study were asked to evaluate the extent to which various factors influenced their vote related to each SSI they consistently ranked political views as the least influential factor. Despite the claimed low influence on their vote, political justifications are the third most popular type of justification used by the students
when supporting their votes with regard to the EPA task, accounting for 14% of the justifications. Students that justified their votes with a politically based argument were usually questioning the role of government entities in the context of the task. The two arguments below offer contrasting views of how this type of justification appeared in students’ arguments. One student suggested:

   Congress needs to make this decision. (Student, T010)

However, another student took a different stance, yet directed toward the same overarching concern about the role of government bodies:

   It is the EPA's responsibility to (as it is stated in the name) protect the environment. It is not Congress's prerogative to choose when they want to step in and take control. The EPA is already a Federal Agency. If its policies are not congruent with those of Congress, it should be disbanded or rechartered. (Student, C009)

   In contrast to the EPA task, the political justifications used to support students’ votes in the context of the Sugar task represented a more personal view of the role of government in regard to individual freedoms. Accounting for 11% of the justifications for this task, many of these justifications expressed a general distaste for government intrusion on the participants’ freedom of choice. Political justifications in the context of the Sugar task seem to take a defensive tone when used to support a vote:

   It is not the government's choice to determine whether or not a person should eat sugary drinks. There is no reason for a blanket tax over everyone when only some have the problem. (Student, C010)

   I would vote against the tax because the government should have no control over what a person buys. If the customer wants to buy soda then they shouldn't be punished. (Student, C011)

   The two least common types of justifications were those that were emotive in nature, meaning those that expressed sympathy or concern for those involved and scientific or empirical justifications, meaning those that referenced specific data or scientific ideas when supporting their vote. Justifications that were emotive in nature were similar to those justifications that were scored as representing the greater good. However, these emotive justifications were more personal or showed some sort of direct concern for those affected based on the context of the task. One particular student expressed his concern for those affected in the context of the Sugar task by stating:
Obesity is a severe problem in America. This tax could help Americans help themselves by making healthy food choices. (Student, C013)

In the context of the EPA task, the concern expressed in emotive justifications was focused on those families that may be directly affected by the potential job losses discussed in the task. The following student was concerned about the possibility of families going without basic necessities:

The country is a crisis right now. So the USA needs all possible type of jobs, so people can supply their families with food and basic stuff. (Student, T013)

Finally, despite being claimed as the most influential factor having an affect on students’ votes, the least common type of justification were those focused on scientific or empirical grounds. In order to be considered a scientific or empirical justification there needed to be some reference to a scientific idea that could help explain or support the student’s vote, or the student could use or reference data included in either task to help support their vote. The use of correct scientific ideas is ultimately an important aspect of students’ justifications and arguments supporting their votes, however, in the context of this study it was only a goal to identify if students used this category of justification as opposed to assessing their use of scientifically accurate ideas. As such, these types of justifications only represented 2% of all the justifications for the EPA task and only 6% of all the justifications for the Sugar task. On average, the most common type of justification used by the students out numbers the scientific or empirical justifications used by a ration of nine-to-one. Some examples of scientific and/or empirical justifications used by the students include:

I would vote in favor of the tax because I support the sin taxes that do not lower availability of the product to the consumer. Data from Arkansas I believe it said, showed that the tax did not change consumption yet it raised 400 million dollars for medicaid. That's enough of a justification for me. (Student, C014)

$95 billion - $16 billion is still $79 billion dollars needed to fund obesity related problems, maybe a better idea is out there. (Student, T014)

These two students both focus on the amount of money that could potentially be raised through the implementation of the sugar tax. The first student uses that data to help

---

6 The actual dollar amount provided in the text associated with the Sugar task is $40 million, not $400 million.
justify voting in favor of the tax, whereas the second student uses similar data to support a vote opposing the tax. In each case the students are referencing specific data that were part of the arguments discussed within the Sugar task. In the context of the EPA task some students were able to use scientific ideas when supporting their vote, consider the following example:

I would vote against the legislation because of the effect it is currently (obviously) having on the environment. The fact is that the coal companies are trying to buy their way out of paying for new technology by claiming that the current uses of coal are clean... they are not. Combustion produces carbon dioxide, carbon monoxide and also other harmful hydrocarbons that were not meant to be in the atmosphere in such high levels. (Student, C014)

This student is able to use what he knows about the combustion process to help justify his vote against the proposed legislation associated with the EPA task. Another student, however, uses his understanding of science differently.

There is no need for the EPA to bypass our elected representatives to destroy our financial well being. There is no clear proof of global warming and every nation that has implemented any legislation like this has paid dearly for it financially. (Student, T011)

This student argues that the scientific implication of global warming which underlies the EPA task is unfounded. In each of these examples there may be concern about the validity of the scientific justifications the students invoke, however, the goal of this study is to simply identify if students attempt to use such means for justifying their votes.

**Differences in justifications between groups**

The design of this study allows for the comparison between the two different groups of students based on the type of instruction they received. The treatment group experienced the ADI-based course described above and the comparison group experienced the more traditional course, also described above. By comparing the nature of the students’ justifications between groups on each SSI task before and after the intervention it is possible to observe any relative changes in the types of justifications the students use that may be related to the type of instruction they received. Figure 3.3 shows how the distributions of justifications compare to one another for each group on the pre-intervention measure for both tasks. The two groups have very similar distributions on each task with some differences between the political and intuitive justifications for the
EPA task (11% difference and 4% difference between groups respectively) and some differences between the intuitive and economic justifications for the Sugar task (10% difference and 7% difference between groups respectively). A Pearson chi square test showed that these differences were statistically different between the two groups in regard to the EPA task, Pearson $\chi^2 (5, N = 131) = 11.39, p = .04$, but not significant for the Sugar task, Pearson $\chi^2 (5, N = 137) = 9.98, p = .08$.

Post-intervention analyses still showed no significant differences between the two groups in regard to the Sugar task, Pearson $\chi^2 (5, N = 110) = 7.22, p = .21$; however, there was a significant difference between the two groups with respect to the EPA task, Pearson $\chi^2 (5, N = 121) = 19.27, p < .01$. Additionally, the significant difference between groups is accompanied by a significant change in justification distribution with the comparison group from the pre- to post-intervention assessments, Pearson $\chi^2 (5, N = 101) = 17.02, p < .01$. The significant difference between the pre- and post-intervention assessment for the comparison group does not necessarily correspond to an improvement in the uses of the various justifications. Comparing the justification distributions in Figures 3.3 and 3.4 indicate, in the context of the EPA task, the comparison group experienced an 11% decrease in greater good justifications and a 10% increase in intuitive justifications, which are the likely cause of the significant difference in distributions, yet such changes could be viewed as a regression in the nature of the argument justification considering justifications based on the greater good show concern beyond the level of the individual and intuitive arguments are those based on a “just because” viewpoint.
Figure 3.3 Comparison of distribution of justifications for the EPA and Sugar tasks between groups, pre-intervention
Post-intervention distribution of justifications and influential factors

There were no observed differences between the groups on the pre- to post-intervention measures regarding the types of justifications the students used to support their votes. During the analysis of pre-intervention arguments some misalignments emerged between what factors the students claimed influenced their vote and the types of justifications they actually used when supporting their votes, these discrepancies are evident when comparing Figure 3.2 and Table 3.3. The distribution of justifications displayed in the graphs of Figure 3.4 show that the students tended to support their votes with arguments focusing on issues related to the ‘greater good’ of society or economic considerations. Similar to the pre-intervention distributions, scientific or empirical justifications were the least used type of support for their vote. Given there were no statistical differences between groups based on their distributions of justifications post-intervention, the types of justifications were aggregated for further comparison to the claimed influential factors. Figure 3.5 shows the distribution of justifications for each task for both groups combined; Table 3.6 displays the relative rankings of the factors claimed to influence the students’ votes.

In the context of the Sugar task, the various factors were ranked with the same relative influence as the pre-intervention. For the EPA task, however, the general science factor was claimed to be more influential post-intervention than pre-intervention, ranking as the second most influential factor for that task. Despite the increase in the influence of general science for the EPA task and the continued high influence of health and nutrition knowledge for the Sugar task, arguments using scientific or empirical justifications were still the least common, accounting for only 4% of all the justifications, down from 8% on the pre-intervention assessment.
Figure 3.4 Comparison of distribution of justifications for the EPA and Sugar tasks between groups, post-intervention
Table 3.5 Mean rating for factors influencing students’ votes, by task, post-intervention

<table>
<thead>
<tr>
<th>Task</th>
<th>Economics</th>
<th>General Science</th>
<th>Morals &amp; Values</th>
<th>Content Specific Science</th>
<th>Politics</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPA</td>
<td>2.80</td>
<td>2.74*&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.67</td>
<td>2.66</td>
<td>1.48*&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(1.15)</td>
<td>(1.14)</td>
<td>(1.32)</td>
<td>(1.22)</td>
<td>(1.44)</td>
</tr>
<tr>
<td>Sugar</td>
<td>2.63</td>
<td>2.34</td>
<td>2.82</td>
<td>3.08</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td>(1.23)</td>
<td>(1.33)</td>
<td>(1.26)</td>
<td>(1.12)</td>
<td>(1.50)</td>
</tr>
</tbody>
</table>

*mean is significantly different from pre-intervention

a. t(151) = 2.25, p = .03
b. t(151) = 2.25, p = .03

Figure 3.5 Distribution of argument justifications for both groups on each task, post-intervention
Structure of arguments supporting students’ votes on SSI

Also of interest in this study was how students structured their arguments when supporting their vote related to each SSI. The structure of the students’ arguments was assessed independently of their justification and each argument was scored on a five-point scale based on the argument’s complexity. Figure 3.6 shows the distribution of argument scores for each group pre-intervention with the scores for both tasks aggregated within each group.

Both the comparison and treatment groups have very similar distributions of argument scores as seen in Figure 3.6. For both groups nearly 50% of the arguments were scored as level one, which indicates the student provided one or more reasons to support his or her vote. The following arguments are representative of a level one argument:

I do support a soft drink tax because it would help our economy expand. (Student, C015)

We need something that is going to help our economy and make this world a healthier state. If this legislation will do that we need to be willing to try this. (Student, C016)

These two students only provide reasons to support their votes, with Student C015 providing one reason related to the economy and Student C016 providing two reasons, one related to the economy and another about the health of the ‘state.’ These arguments were scored as a level one because the students do not expand on why the reasons they have identified are relevant, in other words they have not provided a rationale for their reasons.

Level two arguments, on average, represent 25% of the supporting arguments, these arguments included one or more reasons like those in level one, but in the case of level two arguments a rationale was introduced to help support the students’ vote.

I think that added a sales tax to soft drinks would discourage people from buying said drinks, thus reducing their calorie intake as well as their aptitude for negative health conditions. (Student, C012, emphasis added)

I feel as if it is time to move towards a better source of power other than coal, whether it is clean coal or not. Burning coal pollutes the environment and will not be around forever, one day it will be used up and we will have to resort to alternative sources for power, so why not now? (Student, T008, emphasis added)
The underlined portions of the two arguments above show how some students included a rationale in their argument. Student C012 expands on his notion of decreasing consumption by further explaining why that reduction would be beneficial; Student T008 provides a rationale for her reason of moving toward a new energy source by describing why such a change would be necessary.

Collectively, levels three and four represent about 14% of each groups’ scores. These types of arguments contain a rebuttal to the opposing view or vote; in the case of a level three argument the rebuttal is connected with a list of reasons, whereas with a level four argument the rebuttal is coupled with reasons and a rationale. The sample argument below is from the Sugar task and represents a level four argument.

Obesity is one of the main health issues that Americans face everyday. It affects everyone. By putting a tax on soft drinks, it would discourage more people from buying and consuming the soft drinks, that have a connection with obesity. Also, the money gained from this tax could be used for programs that promote health and proper dieting. **When people say that they are working on a budget, I say that is not even an excuse. Water is much cheaper than any soft drinks, and is sometimes free at some places.** (Student, C017, emphasis added)

In this argument the student provides one reason, discouraging consumption, followed by the underlined rational, due to the connection with obesity, which is followed by a second reason, health programs, and then completes his argument with a rebuttal (in bold) to the notion that the proposed tax would place an undo burden on families with tight budgets. A level three argument is similar in structure, however, there is no rationale included; consider the following example.

If we allow congress to put a ban on EPA it will just allow for more pollution to occur. **Yes if EPA starts shutting down companies because they are polluting too much it could severely hurt the economy, but if we as a society don't start changing the way we pollute we will lose a lot more.** (Student, C018, emphasis added)

This student provides a ‘more pollution’ reason to support her vote, and even though it may be implied, does not include a rationale for that reason. She then goes on to provide a rebuttal (in bold) to one of the claims discussed in the EPA task concerning the economic impacts of the proposed legislation.
Figure 3.6 Distribution of argument scores for both groups, pre-intervention
Finally, a large number of arguments were scored as level zero, 9% and 14% for the comparison and treatment groups respectively; these arguments are simply restatements of the student’s vote or a supporting argument that contradicts the student’s vote. Given the format of the prompts in the assessment, the students’ votes were essentially their claim and they were asked to provide an argument to support that claim. If the claim and supporting argument do not agree then the entire argument lacks internal consistency. Despite the fact that some of these arguments contained reasons and in some cases rationales, they were given the lowest possible score due to the lack of internal consistency. In the following example the student does not seem to be making an argument in support of the tax discussed in the Sugar task, yet the student voted in favor of the proposed tax:

Americans are already broke enough as it is. This world is coming to a disaster. Pretty soon we're going to have to actually pay money to be able to breathe. We already do that in a sense, but it's going to be a direct contact. One's health is their concern and the choice is there's. It brings up that the fact that Medicare or Medicaid pay for half of the doctor bill, but in 50 years when I am that age, those companies will most likely not be around. (Student, T015)

Regarding the EPA task, the following student makes an argument that is aligned with the economic considerations presented as part of the task by the supporters of the legislation, however this student’s vote was in opposition to the legislation:

We have been using coal for many years. While it does pollute the environment it doesn't create a huge impact. Coal is cheap and many Americans, specifically in this time, need to cut back on costs. (Student, T112)

Another student provided the argument:

I would support the ban on regulation. (Student, T016)

In this case the student has simply restated the vote they provided in the context of the EPA task. This ‘argument’ does not attempt to support the student’s vote and as such was scored as a level zero argument.

**Argument structure within each SSI**

The participants in this study provided a vote on a legislative item in the context of two SSI tasks, then the students were asked to provide an argument supporting their vote. The data displayed in Figure 3.6 shows the distribution of argument scores for both
tasks, pre-intervention. These distributions are very similar and mirror each other in the relative abundance of the various types of argument structures. A Pearson chi square test indicated that there were no significant differences between the two groups’ distributions of argument scores on the pre-assessment, Pearson $\chi^2 (4, N = 146) = 6.27, p = .18$, when both tasks were combined.

The groups, however, differed when each task was analyzed separately. The distributions of argument scores for each task are shown in Figure 3.7. For the EPA task there appears to be some differences in the proportions of level two arguments between the comparison and treatment groups. A Pearson chi square test showed that overall the distribution of argument scores between the two groups was indeed different, Pearson $\chi^2 (3, N = 73) = 13.45, p < .01$. Due to the low frequency of level four arguments within the treatment group, level three and level four arguments were condensed into one category. Pair-wise follow-up tests showed that there was only a significant difference when comparing the distribution between level one and level two arguments\(^7\), Pearson $\chi^2 (1, N = 42) = 10.61, p < .01$. Based on this analysis and the visual evidence provided in Figure 3.7, the comparison group generated significantly more level two arguments on the pre-assessment than the treatment group in the context of the EPA task. Specifically, 24% of the arguments generated by the comparison group were level two, while only 8% of the treatment group arguments were level two. It is important to note that level two arguments include reasons as well as a rationale, which is an important structural improvement in the overall sophistication of an argument.

For the Sugar task, the distributions of argument scores between the comparison and treatment groups are almost identical. Given the low frequency of level three and level four arguments within the context of this task, they were once again combined into one level for the purposes of analysis, additionally, level zero arguments were not included in the comparative analysis as only one argument from each group was scored at this level. Based on a Pearson chi square test, there were no significant differences in the distributions of argument scores between the two groups on the pre-assessment for the Sugar task, Pearson $\chi^2 (2, N = 72) = .237, p = .89$.

\(^7\) Level 0 & Level 1 pair-wise follow up comparison: Pearson $\chi^2 (1, N = 55) = .19, p = .66$; Level 2 & Level 3/4 pair-wise follow-up comparison: Pearson $\chi^2 (1, N = 18) = 3.18, p = .07$. 

53
Figure 3.7 Distribution of argument scores for each SSI task, pre-intervention
Figure 3.7 also shows an important difference between the two SSI tasks used in this study; there is a noticeable difference in the amount of level zero arguments between the two tasks. In the context of the EPA task, level zero arguments account for 44% of all the arguments, whereas in the context of the Sugar task, only 2% of the arguments were scored as level zero. Students were much more likely to provide an argument that contradicted their own vote in the context of the EPA task. This is likely due to the increased reading difficulty of the text associated with that task. Based on a Flesch reading ease assessment (Flesch, 1948) the text for the EPA task scored 20.0 and the text for the Sugar task scored 40.1. The scale for the Flesch reading ease test ranges from 0 to 100, with 100 being easier to read and 0 being most difficult to read. According to this scale the text associated with the Sugar task is categorized as “difficult” and the text associated with the EPA task is categorized as “very difficult” (Flesch, 1948). The proposed legislation in the EPA task was also worded in the negative, meaning a vote in favor of the legislation supported the removal or delay in implementing greenhouse gas regulations. This may have been confusing to some students who voted in favor of the legislation thinking they were voting in favor of the regulations.

After the semester long intervention, the participants in this study again completed the two SSI tasks by providing their vote on each issue and a supporting argument. The distributions for the argument scores for each task and each group are shown in Figure 3.8 below. Comparing the post-intervention distribution of scores for the EPA task there were once again significant differences between the two groups; Pearson $\chi^2$ (3, N = 73) = 12.80, $p < .01$, again due to the low frequency of level four arguments, the level three and level four arguments were grouped together for the analysis. Given that the overall distribution of scores was different for the two groups, follow-up pair-wise comparisons were conducted to identify the cause of the significance. The pair-wise comparisons only resulted in a significant difference between the level zero and level one arguments between the two groups, Pearson $\chi^2$ (1, N = 49) = 10.22, $p < .01$. On the post-intervention assessment the comparison group reduced their proportion of level zero arguments by 5% and increased their level one arguments by 18%. The treatment group,

---

8 Level 1 & Level 2 pair-wise follow up comparison: Pearson $\chi^2$ (1, N = 40) = .09, $p = .76$; Level 2 & Level 3/4 pair-wise follow-up comparison: Pearson $\chi^2$ (1, N = 24) = 2.58, $p = .11$. 

55
Figure 3.8 Distribution of argument scores for each SSI task, post-intervention
in contrast, only reduced their level zero arguments by 1% and reduced their level one arguments by 7%. It should be noted that the difference between the comparison and treatment groups with respect to level two arguments pre-intervention, is no longer present in the post-intervention assessment. In other words the discrepancy in the groups’ tendency to craft level two arguments is not present after the intervention. The persistence of level zero arguments, or vote-argument contradictions, in the context of the EPA task indicates that even after the intervention the difficulty of the text may have hindered the students’ ability to generate more sophisticated arguments.

Initially this analysis indicates that the comparison group reduced their proportion of level zero arguments and increased their proportion of level one arguments, this is a positive change, however additional analysis within each group from the pre- to post-assessment may offer more insight to the group changes. When comparing the pre- and post-intervention distributions of argument scores for the treatment group, there were no significant differences in the context of the EPA task, Pearson $\chi^2 (3, N = 73) = 3.00, p = .39$; however, for the comparison group there were significant differences between the two assessments, Pearson $\chi^2 (3, N = 79) = 9.09, p = .03$. Pair-wise follow-up tests show that there are significant differences between the distributions of level zero and level one arguments, Pearson $\chi^2 (1, N = 53) = 4.10, p = .04$ and between level one and level two arguments, Pearson $\chi^2 (1, N = 57) = 4.85, p = .03$. This statistical analysis coupled with a visual analysis of Figures 3.7 and 3.8 tell a different story regarding the changes within the comparison group. The proportion of arguments scored as level one increased by 18%, at the same time the proportion of level zero arguments dropped 5%. This decrease alone, however, does not account for the large increase in the proportion of level one arguments. Even though the proportion of level one arguments increases, the proportion of level two and level four arguments decrease post-intervention. I describe this type of change as a funneling effect where the argument scores seem to funnel toward one value (i.e. level one) as opposed to shifting the distribution.

Applying this same type of analysis to the treatment group, there are no significant differences between the pre- and post-intervention distributions of argument scores, however there is more of a positive shift in argument scores associated with the changes within each level. Level zero, level one, and level three arguments decrease by
approximately 10% and in turn the proportion of level two and level four arguments increase by 4% and 6% respectively. Although these shifts are not large nor are they significant, they represent a positive trend in the progression of the treatment group. These changes represent small shifts toward more sophisticated argument structures as opposed to large, significant shifts toward less sophisticated structures.

In the context of the Sugar task the pre-intervention comparison revealed no significant differences between the distributions of argument scores for the two groups; however, there are significant differences between the two groups when compared post-intervention. A Pearson chi square test shows that there are significant differences in the distribution of argument scores between the two groups, Pearson $\chi^2 (2, N = 72) = 10.86, p < .01$. As in the pre-intervention comparison due to low frequency of scores the level zero arguments were excluded from analysis and the level three and level four arguments were combined into the same level for analysis. Given the overall significant difference in distribution for the two groups follow-up pair-wise comparisons were conducted to determine the source of significance. There was a significant difference between the level one and level two arguments, Pearson $\chi^2 (1, N = 63) = 10.91, p < .01$; however, there was no significant difference between the level two and level three/four arguments, Pearson $\chi^2 (1, N = 42) = 1.18, p = .28$. The distributions for the Sugar task displayed in Figure 3.8 show that the treatment group generated a larger proportion of level two arguments while also generating a smaller proportion of level one arguments with respect to the comparison group. This indicates that after the semester long intervention the treatment group included rationales in their supporting arguments more often than the comparison group in the context of the Sugar task.

Further analysis shows that neither group alone made significant changes between the pre- and post-intervention assessments, however the cumulative effect of small changes within each group resulted in the significant differences between the two groups on the post-intervention assessment. Even though the treatment group increased their proportion of level two and level three arguments by approximately 11% and 7% respectively these changes were not significant compared to their pre-intervention levels, Pearson $\chi^2 (2, N = 72) = 5.63, p = .06$. The comparison group experienced smaller changes in argument score distribution, with no change in the proportion of level one
arguments and only a 3% increase in level three arguments, which resulted in no significant difference from their pre-intervention distribution, Pearson $\chi^2 (2, N = 76) = .45, p = .80$.

**Development of epistemological sophistication**

As discussed above, a student’s level of epistemological sophistication can have an influence on his or her ability to craft an argument. To determine the influence of the students’ epistemological sophistication in this study the students completed an epistemology questionnaire, which was used to categorize individual students as absolutist, multiplist, or evaluativist. The students completed this questionnaire both pre- and post-intervention; the results for the comparison and treatment groups are displayed in Figure 3.9. The distribution of students among the three epistemological categories is very similar for both groups as well as both pre- and post-intervention. Evaluativists represent 27% of the comparison group and 32% of the treatment group, both pre- and post-intervention. Multiplists represent approximately 72% of the comparison group and 68% of the treatment group, both pre- and post-intervention. There are subtle differences and changes in the percentage of students within the absolutist and multiplist categories, however these differences are not significant between groups on the pre-intervention measure, Pearson $\chi^2 (1, N = 72) = 1.06, p = .30$, or post-intervention measure, Pearson $\chi^2 (1, N = 73) = .78, p = .38$; nor are they significant within the comparison group, Pearson $\chi^2 (1, N = 78) < .01, p = .95$, or treatment group, Pearson $\chi^2 (1, N = 72) = .06, p = .80$ (due to the low frequency, absolutists were not included in this analysis).

Within the context of this study, the epistemological sophistication of the students did not shift over the course of one 15-week semester. This is not necessarily an unexpected result, as epistemological commitments are resistant to change and such a short intervention it not likely to result in large shifts (Bendixen and Rule, 2004). Furthermore, there are no shifts in epistemological sophistication in either group, which indicates that both instructional strategies are insufficient at moving students toward an evaluativist perspective within one semester. That is not to say that facilitating shifts in students’ epistemological commitments is not possible, rather the current iteration of these instructional strategies did not engender such results during this study.
Figure 3.9 Epistemological sophistication for the comparison and treatment groups, pre- and post-intervention.
Influence of epistemological sophistication on students’ arguments

Due to the sample size of each group of students it is difficult to perform a rigorous statistical analysis comparing multiplists and evaluativists on each task, based on their argument scores. Disaggregating the two groups based on epistemological categories and then further dividing based on argument score, results in low frequencies of students within each category, which are insufficient for comparison purposes. However, an alternative option is to combine the two tasks, as representing SSI tasks in general, and then compare within each epistemological category across the two groups. It has already been shown above that there are differences in argument scores between the two groups on the two SSI tasks. Thus, the purpose of the following analysis is to identify if there are any changes in argument scores in the general context of SSI tasks as opposed to task specific changes.

An initial analysis was conducted to compare the overall distribution of argument scores by different levels of epistemological sophistication to determine if there was a difference between evaluativists and multiplists. The literature suggests students with greater epistemological sophistication would generate more sophisticated arguments in terms of their structure. In order to make this comparison the two groups of students were combined and the distribution of their argument scores were compared based on their epistemological sophistication as determined by the epistemology questionnaire. Figure 3.10 displays the distribution of argument scores for each epistemological category, which indicates that the two levels of sophistication were similar in their argument scores, except for the proportion of arguments scoring at the highest level. The difference in level four arguments was enough to constitute a significant difference between the two levels of sophistication, Pearson $\chi^2 (4, N = 88) = 18.37, p < .01$. A follow-up analysis was also conducted within both the comparison and treatment groups, resulting in significant differences between the evaluativists and multiplists in both the comparison group, Pearson $\chi^2 (3, N = 42) = 13.07, p < .01$ and treatment group, Pearson $\chi^2 (3, N = 46) = 8.01, p = .046$ (due to low frequencies, level three and level four arguments were combined for the follow-up).
A second level of analysis was conducted to assess the differences in argument scores based on epistemological sophistication across the comparison and treatment groups. When considering the pre-intervention scores of multiplists there were significant differences in the distribution of scores between the two student groups, Pearson $\chi^2 (4, N = 98) = 14.14, p < .01$. Based on the distribution of scores shown in Figure 3.11 the differences in scores seem to suggest that the students in the comparison group crafted more sophisticated arguments with greater frequency, given that 34% of the multiplists arguments were scored at level two while only 16% of the treatment group multiplists scored at level two. Furthermore, 57% of the multiplists in the treatment group generated arguments scoring at level one compared to only 44% within the comparison group. These observations, when taken together, support the notion that multiplists within the
comparison group included a rationale within their argument more often than those in the treatment group on the pre-intervention assessment.

On the post-intervention assessment there were once again significant differences between the two student groups for the argument scores of the multiplists, Pearson $\chi^2$ (4, $N = 100$) = 10.51, $p = .03$. However, this significant difference seems to suggest that the multiplists in the treatment group crafted sophisticated arguments with greater frequency. Within the treatment group multiplists, the number of level two arguments increased by 17% while at the same time experiencing a 19% decrease in level one arguments, which represented a significant change from the pre-intervention assessment, Pearson $\chi^2$ (4, $N = 100$) = 24.75, $p < .01$. The comparison group, however, had a 9% reduction in level two arguments and a 10% increase in level one arguments, although this was not a significant change, Pearson $\chi^2$ (4, $N = 114$) = 5.27, $p = .26$.

With regard to the students classified as evaluativist, there were significant differences in the distribution of argument scores between the two groups on the pre-intervention assessment, Pearson $\chi^2$ (3, $N = 46$) = 18.96, $p < .01$. The source of this difference is likely the large proportion of level two arguments (30%) generated by the treatment group in relation to the comparison group (12%). There are also a larger proportion of level four arguments generated by the evaluativists when compared to the multiplist, 23% of all the evaluativists’ arguments were level four, whereas only 6% of the multiplists’ arguments reach that level. This finding, based on the pre-intervention assessment, is consistent with other research, which indicates students with greater epistemological sophistication are able to produce more sophisticated arguments (Mason and Scirica, 2006; Nussbaum et al., 2008).

On the post-intervention assessment there were no longer significant differences between the evaluativists within each group, Pearson $\chi^2$ (3, $N = 46$) = 3.08, $p = .38$. The distribution of argument scores for each group was very similar after the intervention, and the large difference in level two arguments is no longer present. However, the gap was not closed due to an increase in proportion of level two arguments (increase of 2%) by the comparison group, rather it was due to a decrease (by 10%) in level two arguments by the treatment group. The decrease in level two arguments for the treatment group was accompanied by an increase (by 6%) in the proportion of level three arguments.
Figure 3.11 Distribution of argument scores for students classified as multiplist, pre- and post-intervention.
Figure 3.12 Distribution of argument scores for students classified as evaluativist, pre- and post-intervention
the comparison group also increased in level three arguments (by 7%). Perhaps the most notable difference in the evaluativists’ arguments post-intervention was a reduction by 14% in the proportion of level four arguments generated by the comparison group. The treatment group also experienced a reduction in level four arguments, albeit only by 2%. This decrease resulted in no level four arguments for the comparison group on the post-intervention assessment, which indicates a regression in argument sophistication even though there was a high level of epistemological sophistication.

**Influence of content knowledge on students’ arguments**

Content knowledge may also have an impact on a student’s ability to construct a sophisticated argument, i.e. an argument that includes rationales or rebuttals. Sadler and Fowler (2006), for example, suggest that once students amass enough content knowledge the quality of their arguments may increase, according to their model there are various thresholds of content knowledge that must be crossed before argument quality increases.

In order to assess the impact of content knowledge in the context of this study the participants responded to a series of content questions related to each of the SSI tasks. Comparing the pre-intervention scores between the two groups revealed that the comparison group out performed the treatment group on both the EPA task content assessment, \( t(150) = 2.67, p < .01, d = .43 \), and Sugar task content assessment, \( t(125.6) = 3.11, p < .01, d = .43 \). Given this result, it may be reasonable to assume that the comparison group would have an advantage when generating their arguments because they had higher content knowledge. The analysis below investigates the relationship between content knowledge and the argument scores for each group on the SSI tasks.

To analyze the impact of content knowledge on students’ arguments a median split, based on the scores of all the participants, was used to group the students into a high content knowledge group and a low content knowledge group. The distribution of argument scores for each group, based on level of content knowledge, is displayed in Figure 3.13. Due to the low frequency of level three and level four arguments those levels were combined for this analysis. The distributions displayed in Figure 3.13 indicate that within each group there are significant differences between the students with low content knowledge and high content knowledge. In the case of the comparison group, Pearson \( \chi^2 \)
(3, N = 76) = 12.94, p < .01, there were 11% more level three and four arguments generated by students with high content knowledge than those with low content knowledge.

For the treatment group, level three and four arguments only differed by 2%; however, level two arguments were 11% more likely for students with high content knowledge, Pearson $\chi^2 (3, N = 49) = 9.11, p = .03$. Thus, prior to the intervention, students with high content knowledge generated more sophisticated arguments within each group. Comparing across the two groups indicated that there were no significant differences between the high content knowledge students of each group, Pearson $\chi^2 (3, N = 49) = 2.58, p = .46$, or the low content knowledge students of each group, Pearson $\chi^2 (3, N = 97) = 7.03, p = .07$, despite the overall higher content knowledge of the comparison group.

The post-intervention assessment of content knowledge, once again revealed a significant difference between the comparison and treatment groups, with the comparison group outscoring the treatment group on the EPA task, $t(128.0) = 3.11, p < .01, d = .50$, and the Sugar task, $t(150) = 2.98, p < .01, d = .48$. Considering the post-intervention argument scores for the comparison group, there were no significant changes from the pre-intervention scores for neither the students with low content knowledge, Pearson $\chi^2 (3, N = 70) = 5.55, p = .14$, nor those with high content knowledge, Pearson $\chi^2 (3, N = 88) = 2.54, p = .47$. Additionally, the difference between the argument scores of those students with high content knowledge compared to those with low content knowledge within the comparison group persisted, Pearson $\chi^2 (3, N = 88) = 13.09, p < .01$. (See Figure 3.14 for distributions)

On the other hand, the treatment group showed no significant difference between the distribution of argument scores for students with high content knowledge compared to those with low content knowledge, Pearson $\chi^2 (3, N = 54) = 3.35, p = .34$. The gap between the high and low content groups, within the treatment group, was closed due to a significant change in the distribution of argument scores for the students with low content knowledge, Pearson $\chi^2 (3, N = 92) = 11.82, p < .01$, accompanied by no significant change in distribution of argument scores for the students with high content knowledge, Pearson $\chi^2 (3, N = 54) = 5.83, p = .12$. 

67
Figure 3.13 Distribution of argument scores by content knowledge for both groups, pre-intervention
Figure 3.14 Distribution of argument scores by content knowledge for both groups, post-intervention
The treatment group students with low content knowledge increased their level three and four arguments by 2% from the pre-intervention measures and also increased their level two arguments by 12%, while decreasing the proportion of level zero arguments by 4%. It should also be noted, that counter-intuitively, the evaluativists within the treatment group increased their proportion of level zero arguments by 5% from the pre-intervention assessment.

It was observed that, within the comparison group, there were no significant changes in argument scores from the pre- to post-intervention assessments for the different levels of content knowledge, despite the comparison group outscoring the treatment group on each assessment both pre- and post-intervention. Regardless of the fact that the treatment group had less content knowledge related to the two tasks, they were still able to enhance the sophistication of their arguments, specifically for the students with low content knowledge. Based on this finding, a follow-up analysis was conducted on the distribution of argument scores between the comparison and treatment groups, for the post-intervention assessment. This analysis showed that there were still no significant differences between the groups for the students having high content knowledge, Pearson $\chi^2 (3, N = 54) = 2.07, p = .56$.

However, on the post-intervention assessment, the students with low content knowledge within the treatment group had a significantly different distribution of argument scores than those in the comparison group, Pearson $\chi^2 (3, N = 92) = 23.28, p < .01$. The graph in Figure 3.15 displays the distribution of argument scores for both groups. There is a large difference in the proportions of level one and level two arguments, with the low content knowledge students from the comparison group generating a majority of level one arguments (63%) compared to only 40% by the treatment group. Additionally, the treatment group students with low content knowledge generated 14% more level two arguments relative to the comparison group. Even though the students in the treatment group had less content knowledge, they included rationales in their arguments more often, taking and important step toward more sophisticated arguments, than did the comparison group on the post-intervention measure.
Figure 3.15 Distribution of argument scores for students with low content knowledge by group, post-intervention
CHAPTER FOUR

Conclusions

Limitations

All studies have strengths and limitations associated with them. The design of this study, therefore, also has strengths and limitations associated with it. Consider the structure of the SSI tasks used in this study and their potential impact on the data collected. The narratives were taken directly from materials available on the Internet and kept in their original form with no changes made to their structure or language choice. Keeping the materials in their native format was valued in this project as students will likely not have the luxury of simplified data or information in the ‘real world’ when dealing with similar complex issues, particularly in the context of proposed legislation. This decision may have been particularly influential with regard to the EPA task given the advanced reading level and complexity of its narrative. The inclusion of TV commercials as part of the SSI tasks was also designed to represent a common form of media the students may encounter that presents information related to an SSI, while also helping to make the tasks more engaging. Yet, a downside of the structure of these SSI tasks was the students’ ability, based on personal preference, to forego the narratives provided within each task and only watch the videos. Even though the SSI tasks were purposefully constructed using content readily available to the students in its native format, those decisions may have also influenced the findings of this study.

Another limitation associated with the structure of the SSI tasks used in this study was the positioning of each SSI task in the broader context of a political vote. Each SSI task was related to a proposed piece of legislation and the students were asked if they would vote in support of or in opposition to the legislation. Using such a context may have influenced the extent to which the students engaged with the tasks or the types of arguments and justifications they used. In some instances politics may be purposefully kept out of classroom interactions yet the SSI tasks used in this study required the explicit confrontation of situations that clearly involved political components. The “voting context” may have unintentionally steered the student participants toward a political focus and away from scientific ideas or deeper concerns related to each SSI task. Thus
SSI tasks structured in such a manner may serve to have negative consequences in terms of student engagement and the focus of students’ arguments.

Critics may also find fault with the pre/post-intervention, repeated measures design of this study. The student participants completed the same assessments and SSI tasks pre- and post-intervention. It is likely that students engaged with the activities well on the pre-assessment, however upon realizing during the post-assessment that they had already completed these tasks, their engagement or motivation may have waned. Additionally, there may be concern of a testing effect due to the repeated measures nature of this study. However, given the use of a comparison and treatment group design, any influence due to a testing affect should be accounted for across both groups and therefore present a minimal concern. Along with the potential testing effects one must also consider the circumstances surrounding the assessments. Students in the comparison group were required to leave their classroom to complete the assessments in a nearby computer lab, whereas the treatment group students were able to complete the assessments in their original classroom. The added requirement of relocating for the assessments may have resulted in a lack of effort or motivation for the comparison group students; admittedly this situation was not ideal, but necessary due to the web-based nature of the data collection instruments.

A second limitation related to the data collection process arises when comparing the nature of the intervention with the nature of the data collection activities. This study was designed to assess the extent to which the ADI instructional model enhanced students’ arguments related to SSI. During the intervention the students experiencing the ADI model worked in groups to develop arguments supporting their claim related to the particular investigations they completed. Group discussion and developing ideas as a group is central to the ADI model during data analysis, argument construction, the argumentation session, and peer-review. However, individual students, not groups, completed the assessments used in this study to gage the influence of the ADI model. The individual nature of the assessments adds another layer to the transfer process. Within ADI, transfer of group meanings is essential when the individual students prepare their investigation reports, yet with the assessments in this study the students are required to make sense of data and develop an argument on their own, which typically does not
happen in ADI investigations. Despite this limitation when comparing the data collection process to the structure of the intervention, it is still the opinion of the researcher that the individual nature of the assessments was appropriate. In the context of science literacy, individuals may not always be able to rely on groups when making sense of information for decisions related to a SSI, the ability to synthesize information and develop a reasoned argument in support of your individual stance is essential.

Finally, the quantitative nature of this study provides a robust battery of data, which was used in rigorous statistical analyses. Although these data were carefully collected and analyzed, they only represent one outlet for understanding students’ performance in the context of SSI. A particular area where qualitative measures would have added potentially valuable data was in the assessment of students’ epistemological sophistication. Qualitative data, such as semi-structured interviews, could have provided a more nuanced account of students’ development as compared to a simple questionnaire. The lack of qualitative measures, however, does not diminish the quality of the quantitative data or the trustworthiness of the findings of this study, but the addition of a more qualitative component to this study would have helped to enhance the overall robustness of the conclusions.

In summary, there are three main limitations associated with this study. First, the tasks that were used to measure how students performed on an individual level when dealing with two different complex SSI, may have been difficult for the students to interpret. Second, the specific circumstances surrounding the data collection process for each group may have also influenced the assessment products generated by the students and the amount of time the students engaged with the assessment activities. Last, the quantitative nature of this study provides valuable data related to students’ abilities to generate an argument in these contexts, yet only represents a limited view of the many factors that may contribute to the products generated by the participants. With these limitations in mind, there are five main conclusions that can be drawn from the data collected during this study.
Factors that influence undergraduates’ decisions about socioscientific issues

The students in this study rated content specific science knowledge as the factor that most influenced their vote on the two socioscientific issues followed by morals and values. Literature within the field “suggests that socioscientific decision-making processes are primarily guided by sociomoral factors” and “if this is the case, then science content likely contributes minimally to the actual decisions individuals make in socioscientific contexts” (Sadler and Fowler, 2006, p. 1000). However, considering the results of this study, the students indicated that science knowledge does influence their decision-making and furthermore, based on the content knowledge assessments it would appear that they also have actual science content knowledge to draw upon when making such decisions.

In studies by Sadler and Donnelly (2006) and Sadler and Ziedler (2004c) they found that their participants viewed the SSI as moral issues, suggesting that scientific content knowledge played little role in their decisions relative to the issues, which is contrary to what was found in this study. It should be noted that the nature of the data for this finding is participant self-reported and may have been artificially inflated due to the fact that the intervention and data collection took place during a science class. Alternatively, the participants in this study may have felt that they possessed a sufficient amount of science content knowledge and may have been able to use that knowledge when weighing various perspectives related to the SSI tasks. If in fact the student participants in this study relied on their scientific content knowledge when making their decisions (i.e. their vote), that may imply that they recognized the SSI tasks used in this study as indeed scientific as opposed to being moral or ethical issues.

The relationship between the factors that influence undergraduates’ decisions and the arguments they use to justify their votes

Although the students rated content specific science knowledge as being a major influence on their decisions, the most common type of justification the undergraduates used in their arguments focused on the greater good for society or the environment, representing a moral or value-based justification, and the least common justification focused on scientific or empirical grounds. These finding coupled with the fact that the students scored 60% or greater on both content knowledge assessments pre- and post-
intervention suggests that there is a disconnect between what students rate as being influential and the types of justification they use to influence others. This finding is supported by similar studies by Ziedler et al. (2002) and Sadler (2004a) who found that students do recognize the importance of scientific content knowledge, however they seemed to compartmentalize their science knowledge and personal beliefs when negotiating SSI, with some of their student participants indicating that they would not necessarily consider scientific evidence when trying to persuade someone (Ziedler et al., 2002).

Thus, there are two potential explanations for these observations. First, the lack of an explicit reference to the content knowledge that the undergraduates rated as highly influential in an argument may indicate that such knowledge is not perceived as the most convincing or persuasive type of argument. The students, therefore, might not have focused on such issues when they wrote their arguments because they did not think it would be persuasive to others. Second, the students may not be able to articulate the content knowledge they possess in a way sufficient to make an argument and thus focus on a more utilitarian aspect of their knowledge. For example, their knowledge related to fossil fuel consumption and potential environmental impacts may be distilled down to a vague or generalizable statement, such as ‘pollution is bad,’ or in relation to the Sugar tax, ‘too much sugar is unhealthy.’ Provided that a student based his or her supporting argument on such a ‘sound-bite’ the resulting argument may be perceived by researchers as a moral or value-based argument, which simply supports environmental stewardship or healthy living, when in actuality the student was justifying a stance using his or her scientific content knowledge.

A third, and perhaps less elegant, explanation for the disconnect between the factors the students’ claimed as influential and those they actually used to support their stance could lie within the nature of the data itself. As discussed previously, this data source relies on student self-reports. It is possible that the disconnect between the factors the students claim to be influential and the actual types of justifications they use are a result of students misrepresenting the actual influence of the various factors in their rankings. It is possible that the students claimed scientific ideas had a great influence because they thought that was the right answer or the answer they were supposed to
provide given that they were completing the assessment in a science class. Likewise the students may have downplayed the actual influence of political views due to the perception that politics should not be involved in their classroom activities.

Given the data presented above, it is likely not valid or acceptable to conclude that scientific content knowledge has little role in students’ decision-making regarding SSI. It does seem reasonable, however, to conclude that most students do not explicitly focus on their science content knowledge when providing an argument to support a particular stance on a socioscientific issue. This implies that students are not able to articulate their science content knowledge in a way that they can incorporate it into an argument or they do not perceive science content knowledge as an appropriate way of supporting a decision, perhaps because it is not convincing.

**The influence of epistemological sophistication on socioscientific argument quality**

Epistemological sophistication has been shown to influence students’ abilities to generate arguments (Mason and Scirica, 2006; Nussbaum et al., 2008). Given the influence of epistemology on argumentation, an objective of this study was to investigate if developing the ability to generate a scientific argument also promoted epistemological development. To that end, the students in this study were given an epistemological sophistication questionnaire both pre- and post-intervention. According to the data generated by the student responses to that questionnaire, there were no shifts in students’ epistemological development for either the comparison or treatment group. Given the short nature of this intervention (15 weeks) and how difficult it is to promote and support epistemological change (Bendixen and Rule, 2004), there likely was not enough time for these students’ to shift epistemological commitments.

Despite the fact that there were no shifts in students’ epistemological sophistication there were nonetheless interesting changes in the students’ abilities to generate arguments supporting their stance on a SSI in relation to their level of epistemological development. Initially, there were significant differences between the multiplists and evaluativists within both the comparison and treatment groups. There was also a significant difference between the two groups at both the multiplist and evaluativist level with the comparison group generating more sophisticated arguments. After the
intervention however, there was no longer a significant difference between the multiplist and evaluativists within the treatment group, indicating that the argument-based intervention of ADI helped to close the gap between these two groups of students. There was also no longer a significant difference between the evaluativists of the two groups and the multiplist of the treatment group generated a significantly different proportion of more sophisticated arguments than the comparison group.

These results suggest the ADI intervention promoted the use of rationales for those students considered to be multiplists, thereby increasing the sophistication of their arguments although they had less epistemological sophistication, while at the same time maintaining the quality of the evaluativists’ arguments. This is contrary to the traditional instruction as the evaluativists in the comparison group experienced a reduction in the sophistication of their arguments. Kuhn et al.’s (2000) description of the assertions of a multiplist as “opinions freely chosen by and accountable only to their owners” (p. 311) may help to explain the increase in their use of rationales. If these students recognize that assertions are merely opinions, then they may be more compelled to back-up their opinion in an attempt to convince others that they are correct. After experiencing the ADI model these student may more readily adopt the strategy of a rationale in an effort to make sense of competing opinions, which may indicate some progress toward an evaluativist level of thinking.

The influence of scientific content knowledge on socioscientific argument quality

The students’ understanding of scientific content did not appear to have much of an impact on their ability to generate an argument about a socioscientific issue. As discussed above, the students in this study did not readily use content knowledge in the arguments they wrote to support their stance on each SSI, although they claimed it had a great influence on their decisions. Furthermore, despite having significantly less content knowledge, the treatment group made significant improvements in the quality of their arguments over the comparison group. These findings indicate that the students’ science content knowledge had less to do with their improved arguments about socioscientific issues than their learning of the scientific habits of mind related to argumentation that were promoted by the argument-driven inquiry instructional model. Sadler and his
colleagues (Sadler and Donnelly, 2006a; Sadler and Fowler, 2006b) have described a threshold model of content knowledge transfer, which suggests that students’ arguments do not significantly improve unless they have crossed a ‘threshold’ of science content knowledge. The results from this study, however, indicate that even though students may not gain content knowledge related to a specific SSI, they can still improve in their ability to generate arguments supporting their stance on a SSI. These improvements in argument quality can be attributed to learning ‘what counts’ as a high quality argument, which is a major goal of the ADI instructional model.

**The impact of an instructional model that emphasizes scientific argumentation on undergraduates’ socioscientific argumentation skills**

One of the goals of this study was to test the hypothesis that students who experience an instructional model that emphasizes scientific argumentation would yield higher quality or more sophisticated arguments in a socioscientific context as well. The results of this study, although clearly not definitive, indicate that students in the treatment group, who experienced the ADI instructional model, incorporated rationales in their supporting arguments more so than students in the comparison group, who experienced the traditional lab instruction. This increase was most notable within the Sugar task where there were significant differences between the two groups in the proportion of level one and level two arguments. In the context of the EPA task the treatment group did not perform significantly better; yet they did make improvements in the sophistication in their arguments, whereas the comparison group seemed to funnel toward level one arguments. These results are supported by similar studies that show arguments related to SSI can be enhanced by explicit instruction in argumentation (Zohar and Nemet, 2002). An important difference in this study, however, is that SSI do not play a role in either course, nor was the science content of the course directly related to the SSI tasks the students engaged with, as has been the case in other studies (Zohar and Nemet, 2002; Sadler and Zeidler, 2004c).

With respect to the results presented above, the treatment group seemed to learn how to generate better arguments by participating in a series of labs designed using the argument driven inquiry instructional model and then transferred what they learned to the context of SSI. The improvement of these students was limited to only increases in their
use of rationales as opposed to their use of rebuttals (i.e. level three and level four arguments), which further supports the notion of rebuttals as an advanced argument strategy, which is difficult to employ (Osborne et al., 2004). The lack of improvement in students’ use of rebuttals may be an indirect indicator that the ADI model does not provide ample opportunity to enact this argument strategy. However, I would suggest that the individual ADI activities likely lack this opportunity, as opposed to the instructional model as a whole. If students engage in the ADI activities in such a manner that provides consistent results across the class then there will be little opportunity to rebut another group’s claim or argument. Therefore, if increasing the opportunity for rebuttals is desired, care should be taken when designing activities so that differing claims and arguments may arise from the same investigation.

Implications

Given these results the students within the ADI treatment group became more scientifically literate, than those in the traditional course, in the sense that they were able to appropriate the scientific habits of mind, namely what counts as a quality argument, they learned in the context of an instructional model that emphasized scientific argumentation when generating supporting arguments related to socioscientific issues. However, neither group became more scientifically literate in the sense that they explicitly used their scientific content knowledge when providing an argument supporting their stance on a SSI.

When considering the results of this study, an important factor to consider is the nature of the student participants. This study involved a comparison between students at a four-year university and students at a two-year community college. There is generally a perception that students opting for two-year institutions may be less academically oriented compared to university students given the different entrance requirements for the two institutions. Furthermore, some may view course work at a community college as less rigorous than at a university and moreover, university course work may be privileged by external observers (i.e. potential employers). With such premises in place the community college participants in this study all but had the “deck stacked against them.” Yet despite such ill-conceived notions, the community college students within the ADI
treatment group out performed the university students who experienced the traditional instruction.

The student participants in this study who experienced the ADI instruction were able to craft higher quality arguments related to socioscientific issues; specifically, they included rationales in their arguments more often than the students experiencing the traditional instruction. This result taken alone is noteworthy within the context of scientific literacy. However, when considering the broader impact of this development the students’ progress is more substantial. The inclusion of a rationale within an argument indicates that the students realize that there are other reasons that may be offered in support of similar arguments. By including a rationale the students are further explicating why the reasons they have chosen are valid and acceptable within the context of the task. This development identifies the potential internalization of a scientific habit of mind (i.e. backing claims with reasoned evidence) as well as a useful skill that can be employed in everyday life. If these students see the value of a rationale in strengthening an argument and making ones claims stronger and more persuasive, they will be better served in a variety of everyday contexts which may include their chosen profession, political participation, environmental activism or any number of scenarios.

This study was designed to address a variety of factors, however, additional opportunities still remain to investigate other aspects of scientific literacy within this study. Data was collected to identify students’ conceptions of data and evidence pre- and post-intervention as well as data indicating students’ abilities to identify insufficient data or evidence in the context of the SSI in this study. The future analysis of this data will provide further evidence as to the scientific literacy of the students who participated in this study. Developing scientific notions of what counts as data and evidence as well as the scientific habit of mind of identifying insufficient or missing evidence would indicate student progress toward scientific literacy.

**Implications for college faculty interested in promoting scientific literacy**

If using scientific content knowledge when making decisions in the context of SSI is a desired outcome of science instruction, then such instruction should make explicit the value of using empirical data and evidence in these contexts. Furthermore, it has been
shown that students generally have a poor conception of data and evidence (Sadler, 2004a), which likely impacts their ability to use it when generating arguments. Thus, explicit focus on what counts as quality data and evidence in a scientific context would be beneficial. The ADI instructional model investigated in this study has been shown to help students develop better quality arguments in scientific contexts (Walker et al., 2010) and improve their arguments related to SSI (this study). Therefore, it may be reasonable to assume that structuring ADI investigations around SSI would help students to more readily apply science content in the context of SSI.

Furthermore, the argument-driven inquiry instructional model also provides an opportunity for college faculty to retain a classroom focus on science content (the original intent of ADI) while still promoting the habits of mind students will need when making decisions related to SSI. Incorporating an argument-based instructional model, such as argument-driven inquiry, seems to function as an intermediate step toward teaching for scientific literacy. Most undergraduate courses are structured around a transmission model of learning (Wyckoff, 2001; DeHaan, 2005) or confirmation style activities in the laboratory (Cooper and Kerns, 2006) as a means to address a large body of science content. This ‘traditional’ style of teaching, however, does not serve to enhance students’ arguments in the context of SSI, whereas the argument-based ADI instruction did. Therefore, it is possible for students to progress toward scientific literacy (i.e. making reasoned decisions related to SSI), while still experiencing instruction that privileges science content as opposed to instruction requiring a focus on SSI tasks.

Scientific argument-based instruction may not be the ideal instructional model for achieving the largest gains in students’ abilities to construct well-structured arguments in the context of SSI, however it does consider the very real constraints and goals of college faculty that tend to privilege a classroom focus on science content. Thus, instructional strategies like ADI, are agreeable with teaching goals focused on science content, with the added benefit of developing scientific habits of mind that transfer to SSI contexts. Finally, with the inclusion of ADI activities structured around SSI, it only stands to reason that the students’ abilities to transfer the scientific habits of mind engendered in ADI activities to SSI contexts would be improved.
Implications for those interested in using scientific argumentation to promote better socioscientific argumentation

When considering the nature of the ADI instruction the treatment group received, the course content and focus was not directly related to the SSI tasks the students completed. This intentional feature of the research design was included to position the ADI activities as potential near-transfer situations for the SSI tasks. Given the variety and complexity of SSI and the myriad factors that influence decisions regarding each, the issue of transfer from one situation to another is crucial in SSI related research (Sadler, 2009b). Additionally, a basic assumption of research related to SSI is that the students will actually transfer “knowledge, attitudes, and practices that will better position the student for dealing with other SSI that s/he will confront in the future” (Sadler, 2009b, p. 701).

The results from this study imply that students who experience science instruction focused on developing argument skills in the context of science content were able to transfer their understanding of what counts as a high quality argument in science to a socioscientific context. Despite being able to transfer these basic habits of mind, specifically the use of rationales, the students did not use their science content knowledge to support their viewpoint on an SSI. Students, however, might be more likely to use science content knowledge if the actual content of the course was more aligned with the underlying science content of the SSI tasks. In the case of this research the students seem to have learned how to transfer a particular ‘practice’ from a scientific context to the socioscientific context. While these two situations do not represent a direct transfer they may be similar enough to represent a near-transfer situation in that the students apply similar strategies in both contexts.

Using instructional strategies grounded in scientific argumentation seem like a viable approach to enhancing the structure of students’ arguments in the context of SSI. This improvement in argument structure may be a worthwhile outcome, but it falls short of engendering an inclusion of scientific content knowledge in students’ arguments. Therefore, those interested in using scientific argumentation to promote better socioscientific argumentation should be explicit about the features of scientific argumentation (i.e. reliance on empirical data to support claims) which are also applicable to contexts that are not purely scientific. When it comes to making decisions
related to SSI there are a variety of factors that can influence an individual, the more aware students are that SSI are indeed scientific, at least in part, perhaps the more likely they will be to use their science content knowledge when making decisions.

**Future Research**

At the conclusion of this study there are several interesting and important findings that inform the SSI research community. Additionally, as with all studies, this investigation has lead to several avenues for future research to further inform the community. One of the important findings from this study is that the students claimed scientific content knowledge had a large influence on their decisions related to the SSI tasks. The fact that the content knowledge to which they refer did not appear in their actual supporting arguments indicates that additional research into this area is needed to determine the extent to which students’ espoused reliance on science content knowledge when making decisions related to SSI actually manifests. One particular approach may be a comparison between the students’ rankings of influential factors and actual pieces of data or evidence the students deem most important or influential from the materials they were provided. If the students identify scientific data or evidence as the most influential pieces of information along with claiming science content knowledge as the most influential factor, that would provide evidence to suggest that some other mediating variable is at play when the students generate their supporting arguments (which again, lacked scientific content). Also, a correlation between the actual justifications observed in students’ supporting arguments and the factors they claimed were most influential may provide evidence that students citing science as influential may tend to provide a particular type of justification, which would indicate some sort of re-appropriation of science content knowledge in the context of a supporting argument.

It is also discussed above that, in general, students have a poor conception of data and evidence in scientific contexts. By enhancing students’ conceptions of data and evidence during scientific investigations, particularly by using the ADI instructional model, there may be increases in students’ abilities to use data and evidence appropriately when generating arguments in SSI contexts. This is a claim that could be further explored through additional research that investigates how changes in students’ conceptions of
data and evidence in a scientific context are related to the quality of the arguments they
generate supporting their stance on a socioscientific issue.

Intentional care was taken when designing this study to incorporate media in its
native format related to the socioscientific issues addressed. Providing students with
researcher generated narratives or researcher compiled data sets offer particular
advantages in a research context (i.e. controlling writing style, readability, formatting,
etc.), however, such design strategies insert a layer of abstraction from what students may
actually encounter in their ‘everyday lives’ when dealing with SSI. Rarely will students
be presented with raw data in the media, in the form of news stories or TV commercials,
nor will they likely be able to interpret raw data associated with the many different
scientific domains encompassed by SSI. When students encounter news stories,
commercials and other media sources the data analysis or interpretation, for better or
worse, has already been completed by the authors of the material. Given this set of
circumstances it may be productive to borrow from the field of media literacy when
researching socioscientific issues.

One aspect of the media literacy movement focuses on students’ abilities to
analyze media in an attempt to identify omitted information, targeted audiences, and
potential biases (Hobbs and Frost, 2003). This approach could also be used in the context
of SSI research based on the media presentation of information related to SSI. If students
and the general public are receiving ‘scientific’ information via media sources they need
to have the ability to judge what aspects of those presentations are biased and what may
be missing in terms of supporting evidence and data. Identifying quality data and
evidence in the context of scientific investigations is an integral aspect of ADI; perhaps
this skill is one that could transfer to students’ analyses of media related to SSI, providing
a new avenue for researching socioscientific issues.
APPENDIX A

EPA Task Narrative

Instructions

Below is a press release issued by Senator John Rockefeller (D-WV) describing legislation that he has recently introduced for consideration by the U.S. Senate. Accompanying the proposed legislation are two TV ads, the first ad is from a group called Repower America and the second is from a group called the Clean Coal Coalition. Please read the entire statement and watch both TV ads (you may watch the ads more than once if necessary and you may read/watch in whatever order you choose). Once you have read the proposed legislation and watched the TV ads, please click the link at the bottom of the page to complete and submit the follow-up questionnaire.

Task Narrative

ROCKEFELLER INTRODUCES LEGISLATION TO SUSPEND EPA ACTION AND PROTECT CLEAN COAL STATE ECONOMIES

Washington, D.C.—Senator John D. (Jay) Rockefeller IV today introduced legislation to suspend potential Environmental Protection Agency (EPA) regulation of greenhouse gases from stationary sources for two years.

“Today, we took important action to safeguard jobs, the coal industry, and the entire economy as we move toward clean coal technology,” said Senator Rockefeller. “This legislation will issue a two year suspension on EPA regulation of greenhouse gases from stationary sources—giving Congress the time it needs to address an issue as complicated and expansive as our energy future. Congress, not the EPA, must be the ideal decision-maker on such a challenging issue.

“Two weeks ago, I sent a letter to EPA Administrator Lisa Jackson challenging EPA’s potential regulation of greenhouse gases. Administrator Jackson responded quickly and showed some willingness to move the agency’s timetable for regulation to the end of 2010. This is a positive change and good progress, but I am concerned it may not be enough time. We must set this delay in stone and give Congress enough time to consider a comprehensive energy bill to develop the clean coal technologies we need. At a time when so many people are hurting, we need to put decisions about clean coal and our energy future into the hands of the people and their elected representatives, not a federal environmental agency.”

Congressman Nick Rahall (D-WV) is introducing the House of Representatives companion legislation, with Reps. Alan Mollohan (D-WV) and Rick Boucher (D-VA) as original cosponsors.
Background

Senator Rockefeller has been working to protect West Virginia clean coal and secure the economies in clean coal states. Rockefeller’s legislation will allow two years for Congress to consider comprehensive energy legislation before EPA could begin consideration of regulations. The legislation directs that for two years after enactment the EPA can take no regulatory action and that no stationary source shall be subject to any requirement to obtain a permit or meet a New Source Performance Standard under the Clean Air Act with respect to carbon dioxide or methane, except for the widely-supported motor vehicle emission standards.

The bill will give Congress the time it needs to design and pass well thought-out legislation. Comprehensive energy legislation should be crafted with a combination of certainty and incentives to create the right business atmosphere for coal’s continued use well into the 21st century. In order to give businesses, energy company CEOs, and investors a reason to invest in technology, they need to know there will be a market for that technology and some level of comfort around the certainty of future environmental regulations.

In April 2007, the Supreme Court ruled in Massachusetts v. EPA that EPA must make a determination when it comes to regulating motor vehicle emissions. On December 15, 2009, EPA published its final rule in the Federal Register, stating: “The Administrator finds that greenhouse gases in the atmosphere may reasonably be anticipated both to endanger public health and to endanger public welfare.”

The Supreme Court ruling gives the EPA the authority to regulate greenhouse gases under the Clean Air Act. If Congress wants to change or alter that authority—or suspend it long enough to pass comprehensive legislation—Congress must be able to pass a bill that addresses the real life economic impacts that EPA is not equipped to consider.

To address concerns about EPA’s unmitigated authority in regulating stationary sources and the impact on jobs and local economies, Senator Rockefeller recently led a group of eight Senators from clean coal and manufacturing states in sending a letter to EPA Administrator Lisa Jackson.

The letter conveyed concerns and questions about how the Clean Air Act could be used to regulate a host of greenhouse gas emission sources ranging from coal-burning power plants to factories to oil refineries and many other types of facilities that collectively employ millions of Americans.

Administrator Jackson indicated in her expedient reply that EPA is moving forward with motor vehicle regulations, but the stationary source regulations that are likely to have far-reaching economic consequences will not be acted upon in 2010, but phased-in beginning in 2011.

Senator Rockefeller’s legislation gets to the heart of the matter by providing time for Congress to debate a comprehensive approach to energy policy and its effect on jobs and
our economy without hampering EPA’s ability to move forward with important vehicle efficiency rules.
APPENDIX B

Sugar Task Narrative

Instructions

Some lawmakers in New York have proposed a penny-per-ounce tax on soft-drinks. The following article was obtained from The Center for Science in the Public Interest and discusses implications of such a tax. Alongside the article are two television commercials, one in opposition of the tax, the other in support of the tax. Please read the article and view both commercials (you may view the videos more than once if needed). When you are finished, please select the link at the bottom of the page to complete the follow-up questionnaire.

Why Tax Soft Drinks?

More than two-thirds of Americans are overweight or obese, and soft drinks are the only food or beverage that has been shown to increase the risk of overweight and obesity. Obesity costs $95 billion a year in medical expenditures, of which half are paid through Medicare and Medicaid. Obesity also causes numerous other problems, including reduced self-esteem and lower productivity. Soft drink taxes provide a dual benefit; they could drive down soft drink consumption and generate revenue for much-needed health programs.

Americans Consume Huge Volumes of Soft Drinks

Despite the first-ever per-capita declines in soft drink sales, companies still sold more than 14 billion gallons of calorie-laden soft drinks in 2008. That is equivalent to about 506 12-oz. servings per year, or 1.4 12-oz. servings per day, for every man, woman, and child. Those drinks include regular (non-diet) carbonated sodas, energy drinks, sports drinks, fruit drinks, ready-to-drink teas, and vitamin waters.
Soft Drinks Contribute to Costly Health Problems
Several scientific studies have shown that soft drinks are directly related to weight gain.\textsuperscript{6, 7} Weight gain is a prime risk factor for type 2 diabetes, heart attacks, strokes, and cancer. Frequent consumption of soft drinks is also linked to osteoporosis, tooth decay, and dental erosion.\textsuperscript{8}

A Soft Drink Tax Could Raise Substantial Revenues to Fund Health-Promoting Programs
A federal excise tax of 1 cent per 12-oz serving would raise about $1.5 billion per year. Higher taxes, 4 cents or 8 cents per 12-oz serving, would raise $6 billion and $11 billion, respectively. One cent per ounce would raise more than $16 billion in revenue. The tax would reduce consumption and could fund health-promoting programs, such as media campaigns, encouraging physical activity and healthier diets, providing more fruits & vegetables to school children, healthier school meals, state & local departments of health, biking/hiking paths, and inner-city athletic programs.

Many States Already Tax Soft Drinks
At least 25 states levy special taxes on soft drinks (typically a sales tax).\textsuperscript{9} Some cities also have the authority to levy a soft drink tax. Collectively, they raise well over $1 billion annually. For instance, New York’s tax has been raising several hundred million dollars, and Arkansas has been raising upwards of $40 million. The existing state and city taxes are small and do not significantly reduce consumption. Moreover, little of that revenue is earmarked to promote health (Arkansas has used the revenue to help fund Medicaid).

Time for Action
Health advocates and legislators should work to enact excise taxes on non-diet carbonated sodas, energy drinks, sports drinks, fruit drinks, ready-to-drink teas, and vitamin waters. Those taxes should be large enough to both generate revenues and reduce consumption. The revenue should be earmarked for public health prevention programs, especially in low-income communities.
APPENDIX C

Science Content Assessment

Science Content Questions related to the EPA task

1. Fossil fuels are seen by some scientists as a temporary energy source. Which aspect of fossil fuels encourages scientists to find other sources of energy production?
   a. It takes a very long time for fossil fuels to form.
   b. The formation of fossil fuels results in pollution.
   c. Fossil fuels store energy for a short period of time.
   d. Inadequate organic material exists to replace fossil fuels.

2. Before constructing a new power plant, the builders are required to prepare an environmental impact statement. This document describes the effect that the construction and operation of the power plant will have on the nearby area. Why is it important to prepare an environmental impact statement before building the power plant?
   a. The statement will outline the safety procedures for workers.
   b. The statement helps the community protect is natural resources.
   c. A written statement can be more easily converted into a news article.
   d. The statement will assist the community in planning for new housing developments around the power plant.

3. Which is NOT used as an energy source?
   a. Flowing water
   b. Iron ore
   c. Sun
   d. Oil

4. One advantage of solar energy is that it
   a. Does not pollute
   b. Is not renewable
   c. Is efficient in any climate
   d. Is available at all times

5. What is predicted to be a result of global warming?
   a. Rising ocean level
   b. More severe earthquakes
   c. Larger volcanic eruptions
   d. Thinning ozone layer

6. Three gases found in Earth’s atmosphere are carbon dioxide, nitrogen, and oxygen. What is their order of abundance from greatest to least?
   a. Nitrogen, oxygen, carbon dioxide
   b. Nitrogen, carbon dioxide, oxygen
   c. Oxygen, nitrogen, carbon dioxide
   d. Carbon dioxide, oxygen, nitrogen
7. The burning of fossil fuels has increased the carbon dioxide content of the atmosphere. What is a possible effect that the increased amount of carbon dioxide is likely to have on our plant?
   a. A warmer climate
   b. A cooler climate
   c. Lower relative humidity
   d. More ozone in the atmosphere

8. One of the main causes of acid rain is
   a. Waste from nuclear power plants
   b. Spills from chemical manufacturing plants
   c. Gases from burning fossil fuels
   d. Gases from aerosol spray cans

9. Which group of energy sources are ALL renewable?
   a. Coal, oil, and natural gas
   b. Solar, oil, and geothermal
   c. Wind, solar, and tidal
   d. Natural gas, solar, and tidal

10. Which human activity directly adds the most carbon dioxide to the atmosphere?
    a. Using nitrogenous fertilizer to grow crops.
    b. Burning fossil fuels.
    c. Using air conditioning.
    d. Generating electricity from nuclear reactions.
    e. Increasing the amount of paved surfaces.

11. Oil is an example of a natural resource that is not renewable. Which is another example of a nonrenewable resource?
    a. Wood
    b. Seawater
    c. Sunlight
    d. Coal

---

*Science Content Questions related to the Sugar Task*

1. What is the BEST reason for including fruits and leafy vegetable in a healthy diet?
   a. They have a high water content
   b. They are the best source of protein
   c. They are rich in minerals and vitamins
   d. They are the best source of carbohydrates

2. If the body takes in more food than it uses,
   a. Breathing becomes faster
   b. Weight is lost
   c. The heart rate increases
   d. The food is stored as fat

3. What will most likely affect your adult height?
   a. The height of your parents
   b. The height of your brothers and sisters
c. Your hair color
d. Your weight

4. What are vitamins?
   a. Substances that break down food
   b. Bacteria that people get when they eat some foods
   c. Substances that people make from protein
   d. Substances that people need in small amounts in order for their bodies to function normally

5. The BEST reason for including protein in a healthy diet is because it is the main source of
   a. Energy for the body
   b. Fiber for digestion
   c. Raw materials for cell growth and repair
   d. Vitamins for fighting disease

6. Eating leafy vegetables is important for human health. This is because leafy vegetables are a good source of which of the following?
   a. Protein
   b. Carbohydrates
   c. Minerals
   d. Fat

7. In humans, where does the absorption of food into the blood stream mainly take place?
   a. Stomach
   b. Mouth
   c. Large intestines
   d. Small intestines

8. Which of the following is NOT a component of a sensible weight-loss diet?
   a. moderate calorie restriction
   b. require expensive special foods
   c. behavior modification strategies
   d. adaptable and easy to follow

9. Evidence that there may be a genetic factor influencing obesity is suggested by:
   a. family food habits
   b. peer pressure
   c. exercise habits
   d. identical twins have similar body weights

10. Which of the following is (are) **TRUE** regarding complex carbohydrates?
    a. starches have more Calories than do simple sugars
    b. starch and sugar both provide 4 kcal/gram consumed.
    c. Carbohydrases are made in the liver and secreted into the duodenum via the common bile duct.
    d. Wheat only contains amylepectin, while corn only contains amylase.
    e. B and C are both true.
APPENDIX D

Epistemological Sophistication Assessment

Instructions
Please read the following pairs of sentences and then respond to the question after each pair of sentences.

Robin thinks the government should limit the number of children families are allowed to have to keep the population from getting too big.
Chris thinks families should have as many children as they choose.
1. Which of the following statements do you most agree with about the sentences above?
   a. Only one is right.
   b. Both could have some rightness, but one could be more right.
   c. Both could have some rightness, but one could not be more right than the other.

Robin believes one book's explanation of what atoms are made up of.
Chris believes another book's explanation of what atoms are made up of.
2. Which of the following statements do you most agree with about the sentences above?
   a. Only one is right.
   b. Both could have some rightness, but one could be more right.
   c. Both could have some rightness, but one could not be more right than the other.

Robin has one view of why criminals keep going back to crime.
Chris has a different view of why criminals keep going back to crime.
3. Which of the following statements do you most agree with about the sentences above?
   a. Only one is right.
   b. Both could have some rightness, but one could be more right.
   c. Both could have some rightness, but one could not be more right than the other.

Robin thinks lying is wrong.
Chris thinks lying is permissible in certain situations.
4. Which of the following statements do you most agree with about the sentences above?
   a. Only one is right.
   b. Both could have some rightness, but one could be more right.
   c. Both could have some rightness, but one could not be more right than the other.

Robin agrees with one book's explanation of how children learn language.
Chris agrees with another book's explanation of how children learn language.
5. Which of the following statements do you most agree with about the sentences above?
   a. Only one is right.
   b. Both could have some rightness, but one could be more right.
   c. Both could have some rightness, but one could not be more right than the other.
Robin believes one book's explanation of how the brain works.  
Chris believes another book's explanation of how the brain works. 

6. Which of the following statements do you most agree with about the sentences above? 
   a. Only one is right. 
   b. Both could have some rightness, but one could be more right. 
   c. Both could have some rightness, but one could not be more right than the other. 

Robin thinks one book's explanation of why the Crimean wars began is right. 
Chris thinks another book's explanation of why the Crimean wars began is right. 

7. Which of the following statements do you most agree with about the sentences above? 
   a. Only one is right. 
   b. Both could have some rightness, but one could be more right. 
   c. Both could have some rightness, but one could not be more right than the other. 

Robin thinks people should take responsibility for themselves. 
Chris thinks people should work together to take care of each other. 

8. Which of the following statements do you most agree with about the sentences above? 
   a. Only one is right. 
   b. Both could have some rightness, but one could be more right. 
   c. Both could have some rightness, but one could not be more right than the other. 

Robin believes one mathematician's proof of the math formula is right. 
Chris believes another mathematician's proof of the math formula is right. 

9. Which of the following statements do you most agree with about the sentences above? 
   a. Only one is right. 
   b. Both could have some rightness, but one could be more right. 
   c. Both could have some rightness, but one could not be more right than the other.
APPENDIX E

EPA Task Follow-up Questionnaire

Instructions

Please complete the following questions to the best of your ability.

1. If you, as a citizen, were given the opportunity to vote for or against the legislation proposed by Senator Rockefeller, how would you vote? (select one)

   IN favor of the legislation   NOT in favor of the legislation

2. Provide an argument supporting your choice above regarding the proposed legislation.

3. Not everyone may agree with the decision you have made concerning your vote. Please provide a counter-argument that someone who disagrees with you might use to try and change your mind.

4. How much did each of the following factors influence your decision?

   a. Your religious beliefs
   
      0  1  2  3  4
      Not at all  Some  A great deal

   b. Economic considerations
      0  1  2  3  4
      Not at all  Some  A great deal

   c. Your values or morals
      0  1  2  3  4
      Not at all  Some  A great deal

   d. What you know about science
      0  1  2  3  4
      Not at all  Some  A great deal

   e. The available evidence
      0  1  2  3  4
      Not at all  Some  A great deal

   f. Your political views
      0  1  2  3  4
      Not at all  Some  A great deal

   g. What you know about carbon emissions
      0  1  2  3  4
      Not at all  Some  A great deal

   h. The TV ads
      0  1  2  3  4
      Not at all  Some  A great deal

   i. The written article
      0  1  2  3  4
      Not at all  Some  A great deal
5. What do the terms data and evidence mean to you?

   Data:

   Evidence:

6. What is the most compelling piece of evidence that influenced your decision? Why?

7. Identify two relevant questions, data, evidence, or pieces of information that were not included in the messages or legislation. Why would you like to have that additional information?
APPENDIX F

Sugar Task Follow-up Questionnaire

Instructions

Please complete the following questions to the best of your ability.

1. If you, as a citizen, were given the opportunity to vote for or against the proposed tax on soft-drinks, how would you vote? (select one)

   IN favor of the tax  NOT in favor of the tax

2. Provide an argument supporting your choice above regarding the proposed tax.

3. Not everyone may agree with the decision you have made concerning your vote. Please provide a counter-argument that someone who disagrees with you might use to try and change your mind.

4. How much did each of the following factors influence your decision?

<table>
<thead>
<tr>
<th>Factor</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Your religious beliefs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Economic considerations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Your values or morals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. What you know about science</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. The available evidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Your political views</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. What you know about health &amp; nutrition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h. The TV ads</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i. The written article</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Some</th>
<th>A great deal</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

98
5. What do the terms data and evidence mean to you?

   Data:

   Evidence:

6. What is the most compelling piece of evidence that influenced your decision? Why?

7. Identify two relevant questions, data, evidence, or pieces of information that were not included in the TV ads or description of the proposed tax. Why would you like to have that additional information?
APPENDIX G

Institutional Review Board Approval Letter

Office of the Vice President For Research
Human Subjects Committee
Tallahassee, Florida 32306-2742

APPROVAL MEMORANDUM
Date: 8/16/2010
To: Jonathon Grooms
Address: 4295
Dept.: SCIENCE EDUCATION
From: Thomas L. Jacobson, Chair
Re: Use of Human Subjects in Research
Using Argument-Driven Inquiry to promote scientific literacy and epistemological development

The application that you submitted to this office in regard to the use of human subjects in the proposal referenced above have been reviewed by the Secretary, the Chair, and two members of the Human Subjects Committee. Your project is determined to be Expedited per 45 CFR § 46.110(7) and has been approved by an expedited review process.

The Human Subjects Committee has not evaluated your proposal for scientific merit, except to weigh the risk to the human participants and the aspects of the proposal related to potential risk and benefit. This approval does not replace any departmental or other approvals, which may be required.

If you submitted a proposed consent form with your application, the approved stamped consent form is attached to this approval notice. Only the stamped version of the consent form may be used in recruiting research subjects.

If the project has not been completed by 8/15/2011 you must request a renewal of approval for continuation of the project. As a courtesy, a renewal notice will be sent to you prior to your expiration date; however, it is your responsibility as the Principal Investigator to timely request renewal of your approval from the Committee.

You are advised that any change in protocol for this project must be reviewed and approved by the Committee prior to implementation of the proposed change in the protocol. A protocol change/amendment form is required to be submitted for approval by the Committee. In addition, federal regulations require that the Principal Investigator promptly report, in writing any unanticipated problems or adverse events involving risks to research subjects or others.

By copy of this memorandum, the Chair of your department and/or your major professor
is reminded that he/she is responsible for being informed concerning research projects involving human subjects in the department, and should review protocols as often as needed to insure that the project is being conducted in compliance with our institution and with DHHS regulations.

This institution has an Assurance on file with the Office for Human Research Protection. The Assurance Number is IRB00000446.

Cc: Victor Sampson, Advisor
HSC No. 2010.4570
APPENDIX H

Approved IRB Consent Form

Using Argument-Driven Inquiry to Promote Scientific Literacy and Epistemological Development
Consent Form

You are invited to participate in a research study that is designed to examine the impact of two different instructional models on undergraduate science students’ scientific literacy. You were selected as a possible participant because you are enrolled in an undergraduate science course that requires a laboratory-based experience. We ask that you read this form and ask any questions you may have before agreeing to be in the study. This study is limited to students that are 18 years old or older.

Background Information
The overall goal of this study is to document how undergraduate students’ argumentation skills and ideas about knowledge and learning related to science change over the course of a semester in response to the particular instructional strategy they experience.

Jonathon Grooms is the principal investigator for this study. He is a doctoral student in the School of Teacher Education at Florida State University. Victor Sampson, Ph.D, Assistant Professor of Science Education is supervising this research project.

Procedures
If you choose to participate in this study, you are agreeing to complete two brief tasks designed to allow you to create a scientific argument about a socio-scientific issue and you will complete a brief introductory survey about your understanding of some science topics. These items will be completed once at the beginning of the semester and once at the end of the semester.

Your participation in this study will involve no additional work or time on your part other than these two brief data collection events (approximately 45 minutes each) which will take place during your regularly scheduled class meeting. The information provided by you during this study will be used to evaluate the impact of different instructional strategies on undergraduate science students’ abilities to craft scientific arguments and developments in their thinking processes as a way to promote scientific literacy.

Risks and Benefits of Being in the Study
There is minimal risk, and no direct benefits, from your participation in this study. However, your participation in this study will enable us to improve undergraduate science education.

Confidentiality
The records of this study will be kept private and confidential to the extent permitted by law. In any sort of report we might publish, we will NOT include any information that will make it possible to identify you. Research records will be stored securely and only members of my research team will have access to the records. We will keep the copies of your work stored on a password protected computer. These materials will be destroyed two years after the completion of this research project, which is May 2013.

FSU Human Subjects Committee Approved 8/16/10. Void after 8/15/11 HSC# 2010.4570
Voluntary Nature of the Study

Participation in this study is voluntary and there is no penalty for nonparticipation. Your decision whether or not to participate will not affect your grade in the course or your current and future relations with Tallahassee Community College (TCC) or Florida State University (FSU). If you decide to participate, you are free to withdraw at any time without affecting your grade or those relationships.

In addition, the small class size of this class or the fact that your instructor might be upset with you may make you feel like you have to participate in the study. You do not have to participate if you do not want to. In fact, your instructor will not know if you agreed to participate or not until after he or she submits your grade. Once again, your participation in this study is completely voluntary and choosing to or not to participate in this study will not influence your grade in any way, your instructor’s opinion of you, or your current or future relationship with TCC, and/or FSU.

Contacts and Questions

You may ask any questions you have now. If you have a question later, you are encouraged to contact Jonathon Grooms or Dr. Victor Sampson at:

Jonathon Grooms
Florida State University
Tallahassee, FL 32306-4295

Victor Sampson
School of Teacher Education
Florida State University

If you have any questions or concerns regarding this study and would like to talk to someone other than the researcher(s), you are encouraged to contact the FSU IRB at 2010 Levy Street, Research Building B, Suite 276, Tallahassee, FL 32306-2742, or by phone at [redacted] or by email at [redacted]. You will be given a copy of this information to keep for your records.

Statement of Consent

I have read the above information. I have had the opportunity to ask questions and receive answers about this research project. I consent to participate in the study.

Print Name

Signature

Date

FSU Human Subjects Committee Approved 8/16/10. Void after 8/15/11 HSC# 2010.4570
REFERENCES


BIOGRAPHICAL SKETCH

Jonathon Grooms was born in Charleston, SC and attended The Florida State University (FSU) for his undergraduate studies. In 2004 he graduated from FSU with a Bachelor of Science degree in Secondary Science/Math Teaching with an emphasis in Chemistry and Physics. After earning his B.S. Grooms joined the faculty at FSU as the Director for the physical science outreach program, Science on the Move (SOM), which was a collaborative effort between the Office of Science Teaching Activities and Department of Physics. During Grooms’ six-year tenure as director of SOM he annually worked with over 5,000 K-12 students and over 100 K-12 teachers. In the summer of 2006, Grooms began his doctoral studies in what is now the School of Teacher Education at Florida State University specializing in science education. While pursuing his Ph.D. he has presented research papers at national conferences annually. His research interests lie in the fields of scientific argumentation, socioscientific issues, and instructional models to promote scientific literacy. Grooms will defend his doctoral dissertation in the Spring 2011 semester.