An Exploratory Study on Community College Students’ Reasoning Processes and Argumentation

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Deborah Huei Lan

Graduate Program in Education Department of Teaching and Learning

The Ohio State University

2017

Dissertation Committee:

Dr. Lin Ding, Advisor

Dr. Karen Irving, Co-Advisor

Dr. George Newell, Committee Member
Abstract

The purpose of this study is to explore community college students’ reasoning patterns and argumentation within the context of natural selection. Written assessments and interviews were administered to 2-year college students. Participants were given the Lawson Classroom Test of Scientific Argumentation (LCTSR), a Test of Scientific Argumentation (TSA), and an Open Response Questionnaire (ORQ). The LCTSR and TSA were used to provide insight into student reasoning abilities and students’ understanding of the logical structure of arguments. Significant correlations were found between categories of the LCTSR and TSA, indicating that students’ reasoning and argumentation are connected. The ORQ was used to prove students’ understanding of concepts in natural selection and how they reason through data and coordinate prior knowledge and evidence to develop claims. The interviews were used to analyze how students justify their claims. Results from the ORQ show that students in this sample struggled with identifying, understanding, and using warrants. Additionally, when considering a counterargument and justifying claims, students always set grounds of an argument, rather than critically analyzing the structure or content of counterarguments.
Acknowledgments

I am extremely grateful to my advisors: Dr. Lin Ding and Dr. Karen Irving for providing me support and direction through my program. They were valuable resources for advice on my research interests as well as professional development. Dr. Ding and Dr. Irving were always willing and available to meet and their advice and guidance have been integral to my success in this program. I am also thankful for Dr. George Newell for helping to shape my growth as a student, researcher, and professional. His insights have guided and challenged me to think critically about my own research.

Also, I have special thanks for Dr. Teresa Shiverdecker. Her mentorship allowed me to explore a variety of research interests. Under her leadership, I was given the unique opportunity to find footing both as a researcher and student. Dr. Shiverdecker’s intelligence, kindness, inclusion, and positivity have been inspirational in both my professional and personal life.

This accomplishment could not have been achieved without the help of my wonderful family, who has been so supportive every step of the way. Without the help of my partner, Jason, I would not have been able to complete this journey and I am honored to share this success with him.

Finally, I have special thanks for my colleagues and coworkers: M. Kamb, J. Lemaster, K. Walker, S. Nozaki, and B. Hathaway. Their encouragement has been greatly treasured.
Vita

2008..........................................................Beavercreek High School

2012..........................................................B.S. Biology, The Ohio State University

2016..........................................................M.A. Education, The Ohio State University

2013-2016 ..................................................Graduate Research Associate, Department of Teaching and Learning, The Ohio State University

2016-present .............................................Program Manager, Department of Educational Studies, The Ohio State University

Publications


Fields of Study

Major Field: Education: Teaching and Learning

Program: STEM (Science) Education
# Table of Contents

Abstract .................................................................................................................. ii
Acknowledgments ................................................................................................. iii
Vita ........................................................................................................................ iv
Table of Contents .................................................................................................. vi
List of Tables .......................................................................................................... ix
List of Figures ........................................................................................................ x

Chapter 1: Introduction .......................................................................................... 1
  1.1 Purpose of the Study and Research Questions ............................................. 3
  1.2 Study Overview ............................................................................................ 4
  1.3 Significance of the Study ............................................................................. 5
  1.4 Theoretical Underpinning ........................................................................... 7
    1.4.1 Argument in science. ........................................................................... 7
    1.4.2 Reasoning in science. ........................................................................ 10
    1.4.2 Frameworks for reasoning and arguments. ........................................ 11

Chapter 2: Literature Review ............................................................................... 14
  2.1 Science Education Research – Perspectives and Traditions .................... 14
  2.2 Traditions in Science Education Research Guiding Argumentation Research .............................................................................................. 16
  2.3 Classroom Science and the Philosophy of Science ..................................... 17
    2.3.1 Myths about science........................................................................... 19
    2.3.2 Science and non-science. ................................................................. 21
    2.3.3. Nature of science........................................................................... 21
    2.3.4 Argumentation and classroom science........................................... 22
  2.4 Arguments – Background .......................................................................... 23
    2.4.1 The Toulmin model ........................................................................... 24
    2.4.2 Argument as product. ...................................................................... 25
    2.4.3 Argument as process. ....................................................................... 26
  2.5 Argumentation and the Nature of Science .................................................. 28
  2.6 Arguments as a Social Practice .................................................................... 29
  2.7 The Role of Argumentation in Science Classrooms .................................... 31
  2.8 Evaluating Arguments .................................................................................. 34
    2.8.1 Discourse in argumentation. ............................................................ 34
    2.8.2 Student argumentation .................................................................... 35
  2.9 Reasoning and Explanations ....................................................................... 36
  2.10 Frameworks for Reasoning ........................................................................ 37
  2.11 Community College Students ................................................................... 39
  2.12 Conclusion .................................................................................................. 40

Chapter 3: Methods and Procedures ................................................................... 41
  3.1 Participants .................................................................................................. 42
  3.2 Data Collection and Assessment Instruments ............................................. 45
Appendix C. Open Response Questionnaire ...................................................... 138
Appendix D. Interview Protocol .................................................................... 142
List of Tables

Table 1. Distribution of participants' majors................................................................. 43
Table 2. Participants' ethnicity.................................................................................. 45
Table 3. Number of college level science classes previously taken............................. 45
Table 4. Number of college level math classes previously taken................................. 45
Table 5. Codes for question 1, part a ......................................................................... 59
Table 6. Codes for question 1, part b ......................................................................... 60
Table 7. Codes for question 2, part b ......................................................................... 61
Table 8. Codes for question 2, part a ......................................................................... 62
Table 9. Codes for question 3, part b ......................................................................... 63
Table 10. Codes for question 4, part a ....................................................................... 64
Table 11. Codes for question 4, part b ...................................................................... 65
Table 12. Descriptive statistics: Student scores ....................................................... 68
Table 13. Sub-scales of the LCTSR...................................................................... 70
Table 14. Sub-scales of the TSA.......................................................................... 71
Table 15. LCTSR and TSA Correlations ................................................................. 73
Table 16. Coding and exemplars for question 1, part a ......................................... 76
Table 17. Coding and exemplars for question 1, part b ......................................... 80
Table 18. Coding and exemplars for question 2, part b ......................................... 83
Table 19. Coding and exemplars for question 3, part a ......................................... 86
Table 20. Codes and exemplars for question 4, part a ............................................. 88
Table 21. Coding and exemplars for questions 3 and 4, part b............................... 91
List of Figures

Figure 1. Model for student thinking ................................................................. 12
Figure 2. Argument structure (Toulmin, 1958, p. 97) ...................................... 24
Figure 3. Summary of informal reasoning (from Wu and Tsal, 2007, p. 1167) ........ 38
Figure 4. Open response questionnaire, question 1 .......................................... 51
Figure 5. Open response questionnaire, question 2 .......................................... 52
Figure 6. Open response questionnaire, question 3 .......................................... 53
Figure 7. Open response questionnaire, question 4 .......................................... 54
Figure 8. Interview prompts for question 1 ...................................................... 55
Figure 9. Interview prompts for question 2 ...................................................... 56
Chapter 1: Introduction

This dissertation used a mixed-methods approach to explore students’ argumentation and reasoning processes through written assessments and one-on-one clinical interviews. Rooted in both science education research and the Next Generation Science Standards, scientific argumentation plays an important role in the science classroom (NRC, 2012; NGSS Lead States, 2013). Ohio’s New Learning Standards for Science suggest that scientific inquiry should include processes such as analyzing and interpreting data, thinking critically and logically to connect evidence and explanations, recognizing and analyzing explanations and predictions, and communication scientific procedures and explanations (Ohio Department of Education, 2011). Thus, the nature of science, arguments, explanations, reasoning, and understanding are closely intertwined constructs with complex, abstract, elusive effects on one another (Jimenez-Aleixandre, Rodriguez, & Duschl, 2000; Berland & McNeill, 2010; Cavagnetto, 2010; Bencze & Hodson, 1999). By participating in practices and discourse in the classroom that encourage this type of skill, students are taught that the nature of science requires more than rote memorization to develop scientific knowledge. However, findings in the science classroom indicate that students are not challenged to raise questions or criticisms on the credibility of data (Jimenez-Aleixandre et al., 2000; Nicolaidou, Kyza, Terzian, Hadjichambis, & Kafouris, 2011; Abi-El-Mona & Abd-El-Khalick, 2006). Students treat data as non-biased,
view it as absolute, and use it as links to truths about science, which goes against critical
thinking expectations in the real world (Nicolaidou et al., 2011). Science classrooms need
to shift towards teaching students that knowledge in science is socially mediated. To
translate that into the classroom, teaching science content needs to be done through a
negotiation of knowledge. This negotiation of knowledge depends on consideration of prior
knowledge, experience, content knowledge, and beliefs towards the content topic (Lemke,
1990; McNeill, 2011). Students need to develop scientific skills and practices including
weighing evidence, interpreting texts, and evaluating the viability of scientific claims rather
than viewing science as an uncomplicated compilation of facts that require only shallow
learning strategies to pursue knowledge (Sandoval & Reiser, 2004; Driver, Newton &
Osborne, 2000).

In viewing knowledge acquisition, the nature of science endures as a social practice
influenced by cultural beliefs. Knowledge acquisition occurs through collaborative efforts
and critical analyses of patterns observed in the natural world, models to describe the
patterns, and theories about the underlying mechanistic processes. Developing knowledge
is fluid and dynamic rather than prescriptive, procedural, and stagnant. In authentic science,
scientists are “influenced by factors beyond those internal to science..., factors such as
scientists’ social commitments, values, and by the wider culture of ideas and technological
capabilities in society at the time” (Driver, Newton, & Osborne, 2000, p. 296). Shifting
away from a perspective of “science-as-knowledge” and moving towards a perspective of
“science-as-practice” includes emphasis on scientific practices that cut across disciplines.
These practices include critical analysis of arguments and reasoning, which support
authentic scientific practices by showing that knowing science is not important if there is
no understanding of what a phenomenon is, how it relates to other events, why it is important, and how this particular view of the world came to be (Manz, 2014; Kuhn, 2000, Driver, Newton, & Osborne, 2000; Lizotte, McNeill, & Krajcik, 2004; Christodoulou & Osborne, 2014).

1.1 Purpose of the Study and Research Questions

Despite research occurring along the lines of individualized processes in developing arguments and reasoning, a gap in the literature in looking at the connections between these processes continues to persist. This study sought to explore students’ coordination of personal experiences and beliefs with knowledge from classroom experiences to develop a claim and how they justify their claims when presented with counterarguments. The sample of students had varying experiences in science education. Thus, the content area of the study needed to be general enough such that students could develop a claim despite their level of content knowledge. This study was framed around a complex topic that has multiple perspectives: natural selection. The topic of natural selection has many known student alternative conceptions (Anderson, 2007). Because alternative conceptions have already been explored, the focus of coding and research was on argumentation and reasoning patterns, rather than having to include in depth analysis of content accuracy. Using instruments that probe for students’ understanding of the logical structure of arguments and students’ scientific reasoning skills, connections between those skills and the process for developing arguments were examined. This study is guided by the following research questions:
1) How do community college students perform in reasoning and understanding the parts of an argument and how do these skills correlate?

2) What kinds of claims do community college students use when developing arguments?

3) How do community college students justify and/or defend their claims when presented with a counterargument?

1.2 Study Overview

This study was divided into two parts to explore the research questions: a quantitative/assessment phase and a qualitative/interview phase. In the first phase (quantitative/assessment), students were given a packet that included three assessments: the Lawson Classroom Test of Scientific Reasoning (LCTSR), the Test of Scientific Argumentation (TSA), and an Open Response Questionnaire (ORQ). The LCTSR is a validated, 24 question multiple choice assessment used to explore reasoning skills along the following subscales: conservation of weight, conservation of displaced volume, proportional thinking, identification and control of variables, probabilistic thinking, correlational thinking, and hypothetical-deductive thinking (Lawson, 1978). The TSA is also a validated assessment containing 36 multiple choice items assessing knowledge about argumentation using the following subscales: identifying statements as claim, fact, data, or opinion; identifying qualifiers; labeling statements as claims or non-claims; determining whether support for an argument is based on authority, logic, or theory; and rating the quality of reasoning in an argument (Frey, Ellis, Bulgren, Hare, & Ault, 205). The ORQ consists of 4 open response questions. The first open response question puts forth data from
3 friends fishing and using different fishing pole lengths, fish hook thicknesses, and locations. Participants are asked to identify the connection between those three variables (length, thickness, location) on number of fish caught. The remaining 3 open response questions are question stems from the Conceptual Inventory of Natural Selection (CINS), which is a validated content assessment exploring knowledge about processes related to natural selection (Anderson, Fisher, & Norman, 2002).

To respond to research question 1, correlations between responses to the LCTSR and the TSA were found. This included looking at correlations between subscales of the LCTSR with subscales of the TSA. Research question 2 was explored through coding of the ORQ, which prompted students to state a claim and reasoning behind that claim in question framed in a general, every day setting and three questions framed in the context of natural selection. In the second phase (qualitative/interview), one-on-one interviews with a sub-group of students that volunteered to participate were conducted. These interviews were based on the ORQ with prompts for claim development and justification. Research question 3 was analyzed based on responses to the one-on-one interviews.

1.3 Significance of the Study

Science education research on student thinking explores student argument development and reasoning as a means to understand the processes that underlie student learning and understanding. Argumentation is a process consisting of claim development, evidence identification, and logical justification of the connection between evidence and claims. This process has an internal, individualized component and external, social component. In the internal, individualized component, students are negotiating through
their own experiences and content knowledge to develop an explanation for phenomenon. The external, social component includes justifying an argument against other arguments and critical analysis of others’ explanations. Likewise, reasoning is the coordination of personal experiences, beliefs, and classroom content knowledge when developing claims. This process also has an internal and external component. While the external, social components that accompany these processes have been extensively researched, there is still little understanding of the connections in these processes, particularly in this sample-students in a community college.

This study contributes to science education research in three ways. First, the study contributes to understanding of student reasoning and developing claims. While argumentation and reasoning are related processes, skills across each construct are not extensively explored. This study aims to find connections between specific categories of reasoning and argumentation to investigate whether student reasoning abilities may impact their argumentation. The interview analysis provides additional insight into patterns in reasoning that are associated with developing claims and the patterns of reasoning that are associated with justifying arguments. By examining the types of claims students develop in specific contexts and the reasoning that connects their knowledge to that claim, a deeper understanding of students’ reliance on personal experiences and classroom content knowledge can be developed. Understanding the significance of personal experiences or classroom content knowledge on how students develop claims contributes to research on student conceptual understanding. Second, this study provides a deeper view of how students reason and identify parts of arguments, including more knowledge about the influence reasoning skills has on students’ arguments, providing depth beyond
argumentation and reasoning as discrete units. Third, this study explores the persuasive component of arguments by analyzing counterarguments and rebuttals and comparing whether there are differences in ways students consider arguments based on the context of the question. For example, whether students analyze counterarguments differently based on whether the question is presented within an everyday context, such as 3 friends fishing or presented within a content area, such as origin of variation. In the persuasive component of arguments, this study analyzes whether students’ persuasive techniques for defending their arguments are impacted by context and how resistant their claims are to change.

1.4 Theoretical Underpinning

This study views learning as a social-cognitive process, driven by Vygotskian theory which describes learning as a social process, taking place through interactions with environments (Vygotsky, 1997). The social constructivist notion for knowledge construction aligns with that of science practices. Development of science knowledge is not an individual task but rather a group effort. This view of learning connects well to argumentation because the discourse involved in engaging in argumentation draws social constructivism.

1.4.1 Argument in science.

Argument in science apply general argument structures, such as claim, data, and warrant, to science contexts. In science, claims in arguments are explanations for phenomena observed in nature and data are observations of phenomena. A warrant provides the reasoning for the connection from evidence to claims, usually built upon
existing science principles. While early studies focused primarily on the Toulmin model and its domain general structure for a logical argument as a means to analyze student written arguments, further research provides a critical analysis of how components of data apply to science looking more at the practice of developing an argument and how that discourse builds science knowledge (Cavagnetto, 2010; Wolfe, 2011; Manz, 2014). These studies began to shift away from solely analyzing the occurrence of parts in the structure of an argument and towards how the components of the structure arise and are specific to science. Findings associated with this line of research are that evidence and warrants in scientific arguments are points of weakness (Bencze & Hodson, 1999; Sandoval & Millwood, 2005; Osborne, Simon, Christodoulou, Howell-Richardson, & Richardson, 2013). Students view evidence as incontrovertible and irrefutable, guiding to one specific conclusion (Anderson, 2007). However, in scientific practice, developing a conclusion or explanation for patterns in data often require multiple iterations of analysis and, eventually, making a decision by weighing the viability of many possible explanations. In this vein, there are two main ways to study arguments: multiple explanations and persuasion (e.g. Bencze & Hodson, 1999; Sandoval & Reiser, 2004; Osborne et al., 2013; Osborne, Henderson, MacPherson, Szu, Wild, & Yao, 2016). This includes how students respond to rebuttals and counterarguments and persuade others of their own explanations. By participating in argument development and consideration of multiple explanations, students expand their understanding of scientific practices as well as strengthen their content knowledge (Sadler & Zeidler, 2005; McNeill, 2011). Beyond structural analyses of scientific argumentation are studies that view arguments in science as rhetorical, persuasive, and contextualized (Sandoval & Reiser, 2004 Wolfe, 2011; Osborne et al.,
These studies look at students’ arguments within socio-scientific contexts and science contexts. In particular, how the process of reasoning to develop arguments can impact student content knowledge.

Frameworks for analyzing arguments are guided by the notion that science is more than a body of knowledge (e.g., Benze & Hodson, 1999; Osborne, Erduran, & Simon, 2004; Sandoval & Reiser, 2004; Khishfe, 2012; Driver, Newton, & Osborne, 2000) and support for argumentation discourse views it as a means to engage students in scientific practices including ideas about science learners and science learning.

Scientific argumentation can also be viewed as a dialectic process involving both construction and critique- a competence that is a “complex process of reasoning utilized in situations that require scientific content knowledge to construct and/or critique proposed links between claims and evidence” (Osborne et al., 2016, p. 825). Osborne et al. (2016) uses a framework that looks at arguments as having a critical component that requires justifying and refuting evidence. This cognitive process incorporates persuasion and internal analysis of flaws of arguments.

This study views scientific argumentation as a process that can include an individual, rhetorical component as well as a social, dialogic component. Both of these components include negotiation of meaning around structural parts of arguments as well as constructing meaning for these structures. In argumentation, contextualizing responses is an important step in the process (Berland & Hammer, 2012). Additionally, argumentation contains a persuasive piece that involves justifying the product, which is an argument. Individualized argumentation is viewed as a process that leads to an argument. Through social, dialogic argumentation, participants justify, defend, and provide reasons
for their arguments, creating an explanation. Explanations are arguments that have gone through iterative cycles of argumentation until a socially constructed, negotiated product is formed.

Scientific argumentation consists of both the individual and social processes of argumentation but framed within science contexts. For example, scientific argumentation can include individual considerations on evidence and how evidence is defined in science or social dialogues on weighing evidence and determining parameters for classifying the soundness, accuracy, or acceptability of evidence. Another part of discourse in the science classroom involves scientific reasoning.

1.4.2 Reasoning in science.

In science, reasoning is a combination of personal experiences and classroom knowledge, such as understanding of science principles used to express justification for a claim based on evidence. Lawson (1978, 2000) places the thinking associated with reasoning in science into six categories: conservation of weight and volume, proportional thinking, identification and control of variables, probabilistic thinking, correlational thinking, and combinatorial thinking. In addition to these categories of student thinking and reasoning, science reasoning has been generally described as a students’ coordination of claims and evidence. Like argumentation reasoning can also contain an individual and social component. Individual reasoning can be broadly categorized by three patterns: emotive, rationalistic, and intuitive. The social component of reasoning includes justifying claims and providing context and grounds to defend arguments. Reasoning has been described as a process closely associated with developing arguments and theories (Osborne
Reasoning can be viewed as either an implicit or explicit process, inductive or deductive, informal or formal. This study focuses on reasoning as justification for arguments. Embedded within this process is coordinating prior experiences with classroom knowledge and looking at how they connect.

### 1.4.2 Frameworks for reasoning and arguments.

While there are a variety of frameworks for understanding reasoning and argumentation separately, few explore the implicit process of reasoning and assessing the end product as an argument. This study focuses on informal reasoning patterns characterized as rationalistic, emotive, and intuitive (Sadler & Zeidler, 2005). Rationalistic reasoning is pragmatic. This pattern of reasoning considers multiple variables in the context. By comparison, emotive reasoning is guided by emotion. For example, in the context of natural selection, this would be exemplified as belief or faith-driven thinking. Intuitive reasoning is the “gut feeling” where students are drawn to a particular claim without any explicit support. Student responses can fall in one of these three general reasoning patterns. The product of student reasoning is an argument. Often times, this consists solely of a claim. More complex reasoning patterns that utilize combinations of rationalistic, intuitive, or emotive patterns can lead to stronger arguments that have solid cohesion between parts and are logically sound (Sadler & Zeidler, 2005). In social argumentation, students use critical analysis of arguments to evaluate others’ arguments and justify their own (Abi-El-Mona & Abd-El-Khalick, 2006; Berland & Reiser, 2011). When students put forth a claim as a means to respond to a question, the student is internally negotiating through any number of claims based on those prior experiences, beliefs, and
knowledge (Duschl, 2008). This process can be impacted by many variables such as content knowledge, student reasoning capabilities, and understanding of the soundness of arguments. Osborne et al. (2016) developed a learning progression for argumentation, pointing towards persuasion as a goal of arguments. Guided by this conceptualization of the purpose of arguments, the eventual product through reasoning and persuasion is an argument. Figure 1 shows a visual representation of the iterative process describing students’ thinking based on frameworks by Wu and Tsai (2007) and Sadler and Zeidler (2005).

Figure 1. Model for student thinking
This model shows the connections between components of the study. The ORQ elicited the process of informal reasoning, probing the nuances in rationalistic, intuitive, and emotive reasoning patterns. The LCTSR and the TSA explored additional variables that impact students’ responses to the ORQ. Interview data was used to understand more about students’ formal reasoning. The focus of the interviews was how students justify their arguments when challenged and the persuasive patterns used to convince others of their claims. As students’ scientific thinking develops, the cognitive control of coordinating evidence and prior knowledge becomes more controlled and intentional advancing to a state of meta-cognition about the process. In this sense, a student with the highest level of scientific thinking will be able to not only proceed through the process, but also have meta-awareness of the progression (Kuhn & Pearsall, 2000).

The next chapter provides a brief review of literature on argumentation, reasoning, and classroom science. The review of literature describes a variety of perspectives and frameworks for analyzing argumentation and reasoning followed by a discussion of how those frameworks guided this research study.
Chapter 2: Literature Review

This literature review will first discuss perspectives in science education research, giving an overview for the frameworks of analyzing arguments that fit in each perspective. Following that is a review on the philosophy and nature of science and how science is taught in the classroom, which will frame a context within which scientific argumentation fits showing that its place is not only in the field of science education but also in scientific practice. Next will be a description of literature on arguments including a discussion of reasoning, explanations, arguments and various methods for analyzing reasoning. The last section of this literature review describes the uniqueness in the population of students focused on for this research study. At the end of the literature review is a summary of the chapter and its application to the study.

2.1 Science Education Research – Perspectives and Traditions

Anderson (2007) describes three main traditions in science education research conducted through five broad perspectives. Science education research concerns itself with two core questions 1) Why don’t students learn what we are trying to teach them? And 2) Why does the achievement gap persist? There are five broad perspectives that are explicitly or implicitly addressed in all research studies on science learning:
• Intellectual history and related disciplines (intellectual roots from history or related disciplines). Scientific argumentation analysis utilizes the Toulmin model, which is rooted in disciplines of law and philosophy. This perspective considers changes in the field of science and how this domain general model can be a frame to understand science (Toulmin, 1958; Wolfe, 2011; Manz, 2014).

• Ideas about the nature of science (methods of developing new knowledge, believing that science contains within it specialized languages, values, and practices). In this perspective, argumentation is regarded as process that can engage students in science practices including critical analysis of evidences, justifying claims, and discourse. This is guided by the notion that science is more than a body of knowledge (e.g., Benze & Hodson, 1999; Osborne, Erduran, & Simon, 2004; Sandoval & Reiser, 2004; Khishfe, 2012; Driver, Newton & Osborne, 2000).

• Ideas about science learners and science learning (gaining understandings about the backgrounds, knowledge, beliefs, cultural practices, and other factors that impact a science learner in science learning). Argumentation supports ideas about science and science learning by providing a method of analysis for students’ knowledge, beliefs, and prior experience to negotiate socially constructed meaning (e.g., Tippett, 2009; Jimenez-Aleixandre, Rodriguez & Duschl, 2000; Sandoval & Millwood, 2005; Wu & Tsai, 2007).

• Research goals and methods (scrutiny of qualitative and quantitative methods in search of best practices for conducting educational research). These perspectives are concerned with ways to analyze argumentation. These ways include structural
analyses with quantitative methods by designing large scale surveys for analyzing arguments or qualitative methods of discourse analysis.

- Ideas for improving science learning (focusing on the shortcomings of formal science and ideas about how schools and science teaching could be changed so that students learn more successfully). In this perspective, the focus of research is on argumentation and its link to content knowledge and students’ conceptions of the nature of science.

### 2.2 Traditions in Science Education Research Guiding Argumentation Research

There are three guiding traditions in education research: conceptual change, sociocultural, and critical (Anderson, 2007). In education research, the conceptual change tradition traces back to Piagetian developmental research, investigating students’ understanding based on interactions with nature and the world. Underpinning this tradition is the idea that learning occurs when there are instances of conflict where students can see the contrast between science content and their alternative conceptions. This tradition includes research that falls under the category of exploring student learning, thinking, concepts, and epistemology. There is also a specific field of research that focuses on students conceptual understanding that includes studying students’ mental models and alternative conceptions. This subtle difference is important to note, but will not be a part of the focus for this study. Research on argumentation and reasoning within the conceptual change tradition includes teacher questioning techniques that prompt students to consider multiple explanations for a single phenomenon, student reasoning to create an argument or explanation for a phenomenon, and the structure, quality, and cohesion of students’
dialogic or written arguments (Lee, Liu, Pallant, Roohr, Pryputniewicz, & Buck, 2014; Berland & Hammer, 2012; Chin, 2007; Simon, Erduran, & Osborne, 2006).

In the sociocultural tradition, research is guided by Vygotskian notions of gaining knowledge – that learning occurs from participating in activities with other people. Sociocultural researchers view knowledge building in science as discursive activity. Learning requires understanding of multiple levels of discourse. Research on argumentation in the sociocultural tradition focuses on discursive actions between teacher and student or between students in collaborative argumentation, within a framework that emphasizes the social constructionist view of knowledge (Cavagnetto, 2010; Abi-El-Mona & Abd-El-Khalick, 2006; Jimenez-Aleixandre, Rodriguez, & Duschl, 2000; Chin & Osborne, 2010; Evagorou & Osborne, 2013).

Lastly, the critical tradition focuses on hidden conflicts, both conceptual and cultural in science learning, concerning itself with the ways in which these conflicts are shaped and how their outcomes are determined by power and ideology. Critical researchers view science as inherently ideological and institutional with bias. There are few studies on arguments from a critical tradition standpoint. The focus of research on argumentation is in its connections to learning in terms of conceptual understanding and incorporating the social components of scientific practice.

2.3 Classroom Science and the Philosophy of Science

Science education studies from the 1960s did not distinguish between the process of doing science and content of science, citing that knowledge produced by science cannot be separated from the social practices of its production (Ryu & Sandoval, 2012). Epistemic
practices of science characterize science as a socially and culturally embedded process in which members of the community propose, justify, evaluate, and legitimize knowledge claims within a disciplinary framework. In science learning, there needs to be a move away from declarative knowledge and towards an understanding that knowledge construction should be dialectic between construction and critique (Christodoulou & Osborne, 2012). Students have limited understanding of the epistemology of science. Their understanding about the nature of science tends to be novice and flawed, based on colloquial meanings of terms also used in science (Driver, Leach, Millar, & Scott, 1996). The shift away from what we know to how we know and why we believe in the science classroom has provided a foundation for cognitive scientific processes, such as argumentation, to flourish. Research has shown that having a better understanding on how scientific knowledge is constructed makes one better at doing and learning science (Duschl, 2008).

Sfard (1998) developed two metaphors for types of learning: the acquisition metaphor and the participation metaphor. The acquisition metaphor views knowledge as a “good” that can be attained and accumulated, citing that the “human mind is a container to be filled with certain materials and the learner as becoming an owner of these materials” (Sfard, 1998, p.5). Once this knowledge is attained, it can be transferred to different contexts and shared with others. Students’ views of science practices align with the acquisition metaphor. Teachers have accumulated knowledge about science and transfer the knowledge to students. In the participation metaphor, learning is viewed as an apprenticeship in thinking. There is less focus on the state of knowing and more attention to activities. By comparison, the “learner should be viewed as a person interested in participation in certain kinds of activities rather than accumulating private possessions”
surrendering to the flux of doing rather than the permanence of having (Sfard, 1998, p. 6). By comparison, the practice of science focuses on the act of doing authentic science. Knowledge is built through social construction and what is known in science and about science is in a constant state of flux. The idea of learning as a cognitive apprenticeship has paved the way for a variety of research concepts towards understanding students’ epistemic beliefs about knowledge.

The nature and philosophy of science is important to consider when studying reasoning and argumentation because practices in argumentation such as deliberating on multiple explanation, analyzing evidence, and justifying claims align with science practices. To encourage students to engage in argumentation, a better understanding of the nature of science has to occur (Kuhn & Reiser, 2006).

### 2.3.1 Myths about science.

Much like the metaphors of learning, classroom science has seen similar shifts as the philosophy of science and nature of science increase their roles in classroom learning to close the disconnect between the positivist view of teaching science and the social construction methods of science practice. The differences, as discussed by Driver, Newton, and Osborne (2000) are most accurately expressed: “Science in schools is commonly portrayed from a ‘positivist perspective’ as a subject in which there are clear ‘right answers’ and where data lead uncontroversially to agreed conclusions” (p. 288). By comparison, scientific practices “such as assessing alternatives, weighing evidence, interpreting texts, and evaluating the potential viability of scientific claims are all seen as essential components in constructing scientific arguments” (p. 288). In essence, science is
taught in the classroom as an individual activity, including definitive concepts, responses, and claims that are either correct or incorrect. Even when taught the “scientific method,” students are given prescriptive, step-by-step processes that require very little collaboration amongst each other (Driver, et al., 2000). However, real world construction of knowledge in the sciences is largely socially mediated. Collaborations play a significant role alongside investigations into the natural world that help scientists construct models to explain our perceptions of reality. Traditionally, views of science and the way that it was taught in the classroom had an underlying positivist theme. Similarly, and in line with the acquisition metaphor, science has been taught with the notion that there is a reality about the natural world that can be probed through science experiments until a universal truth is discovered. These myths continue to persist. Students believe there is a hierarchical structure and process to obtaining scientific knowledge that includes a linear scientific method used to prove the truth of hypotheses which turn into theories and eventually laws (McComas, 2006).

In the practice of science, laws describe principles and patterns in nature. Theories are explanations of patterns in nature. Understanding of the relationship between laws and theories leads to an understanding of the tentative nature of science. Science does not seek to prove an absolute truth. Rather than being a static body of collected truths, the practice of science is dynamic, fluid, and tentative because it involves explorations into the real world. Science is not an accumulation of facts. Classroom science leads students through “cookbook” experiments that are procedural rather than exploratory and classroom teaching often inadvertently misleads students into the alternative conceptions about the nature of science (McComas, 2006; Karakas, 2009; Schwartz & Lederman, 2002).
2.3.2 Science and non-science.

In a seminal work, Sir Karl Popper wrote about distinguishing “science from pseudo-science” (Popper, 1999). He identifies that science includes one central criterion: falsifiability. In science, repeated tests and observations are made in an attempt to refute laws and theories. Pseudo-science seeks confirmatory evidence in support of a theory. This method of looking only for evidence to indicate truth is an example of one common student misconception about classroom science.

Understanding the nature of science and identifying what qualifies as science as opposed to “non-science” is often difficult to translate from science practice to classroom practice (Mellado, Bermejo, Blanco, & Ruiz, 2008; Waters-Adams, 2006). In-service teachers tend to rely on their experiences with science as a default for teaching, often elaborating the misconceptions that accompany science learning, without realizing it (Schwartz & Lederman, 2002; Karakas, 2009).

2.3.3. Nature of science.

The nature of science is classified by five key tenants: the social and cultural nature of science, subjective nature of science, empirical nature of science, tentative nature of science, and creative nature of science (Lederman, Abd-El-Khalick, Bell, Schwartz, 2002; Schwartz & Lederman 2002). The social and cultural nature of science views the practice of science as one that is developed socially and is impacted by cultural factors which lead to the subjective nature of science. This tenant of science allows for multiple interpretations of data and multiple theories to explain phenomena observed in nature. However, the nature of science is not completely objective because inherent biases will always exist. The
empirical nature of science describes knowledge development as an iterative cycle of making observations and inferences and looking for patterns rather than a linear process. Naïve views believe that a specific method needs to be implemented to collect facts that prove theories to be true. In fact, there is imagination and creativity involved in science practices because there is no specific cookbook method to follow in order to expand scientific knowledge. Finally, the tentative nature of science is a culmination of the other core concepts of the nature of science. Because science does not seek to prove and observations and inferences can constantly be made about the natural world, there is an inherent fluidity.

2.3.4 Argumentation and classroom science.

In practice, efforts to shift classroom science teaching towards authentic practices in science and away from prevailing misconceptions have driven standards based frameworks to align content to processes and methods in mastering science practice rather than rote memorization of content topics. The National Research Council Framework for K-12 Science Education discusses three dimensions around which science education in grades K-12 should be built. These three dimensions include scientific engineering practices, crosscutting concepts, and core ideas. Included within these essential science and engineering practices is engaging in argument from evidence (NRC, 2012). Research has shown that argumentation is a method that helps students approach a deeper conceptual understanding which, in turn, enables students to produce a higher quality of argumentation (Jimenez-Aleixandre, Rodriguez, & Duschl, 2000). Argumentation also plays an important role in scientific inquiry (Driver, Newton, & Osborne, 2000; Jimenez-Aleixandre, &

2.4 Arguments – Background

Classical theories of arguments use three categories to describe the three main types of argumentative appeals: logos, pathos, and ethos (Wolfe, 2011). Logos is the wording of the argument itself and its underlying logic. Pathos is an appeal to the emotions and ethos assesses the reputation or credibility of the speaker. Beyond argumentative appeals, argumentation theory encompasses multiple facets including analytical, dialectical, and rhetorical schemes for communicating knowledge (Jimenez-Aleixandre, Rodriguez, & Duschl, 2000). Analytical arguments draw a conclusion from a set of premises either inductively or deductively. Dialectical arguments are dialogic conversations with traits that are characteristic of discussion or debate, stressing knowledge and persuasion. These arguments concern themselves more with persuasion compared to previous forms of argument that look more to evidence. These schemes lay the groundwork for science education research seeking to explicate the logic underlying scientific arguments, the context for scientific arguments within society, and critical analysis of how an argument was constructed. Additionally, science researchers looked at methods for reviewing analytical, dialectical, and rhetorical arguments within the framing of science education. This began with identifying the problematic aspects of the science classroom and looking at how argumentation could help to alleviate those problematic aspects.
2.4.1 The Toulmin model.

In previous studies, researchers often used the Toulmin model to evaluate the logical quality of student arguments. The Toulmin model identifies key elements of arguments. It is usually a very powerful tool to evaluate students’ arguments. He found that some components of arguments are field-invariant. However, there are specific definitions for what qualifies as acceptable in each component that are field-specific. Rooted in ideas of philosophy and law, Toulmin created this model to show connections between claims, data, warrants, backings, and qualifiers. This model, found in Figure 2, relates the data (D), qualifier (Q), claim (C), warrant (W), backing (B) and rebuttal (R). There is a logical flow from claim to data on the basis of warrants, supported by backings. These claims are only found to be true under particular circumstances (qualifier) with consideration of rebuttals. Figure 2 provides a visual representation of the Toulmin model.

![Toulmin Model Diagram]

*Figure 2. Argument structure (Toulmin, 1958, p. 97)*

To understand arguments, one must critically analyze the connection between the premise and the conclusion and make judgments on whether that connection is valid.
Current studies in argumentation have aligned with the conceptual change tradition and the sociocultural tradition of conducting research on science learning, citing the five aspects of science that are addressed in research studies.

### 2.4.2 Argument as product.

Studies that focus on the structure of arguments focus primarily on argument as product rather than the process through which arguments are created (Berland & McNeill, 2010; Lee et al., 2014; Manz, 2014; McNeill, 2011; Simon, Erduran, & Osborne, 2006). These studies use the Toulmin model to analyze student production of arguments based on how the claims are supported, whether the argument includes rebuttals, the appropriateness of the claim, and/or the appropriateness and sufficiency of the information included in defense of the claim (Berland & McNeill, 2010; Lee et al., 2014). The frameworks for these studies describe how to discriminate between novice arguments and expert arguments, either by designing learning progressions or proving criteria for what counts as evidence, claims, or warrants and how these benchmarks are specific to scientific argumentation.

Other studies utilize multiple theories as a means to engage students in argumentation based on the idea that thinking and arguing cannot be separated (Osborne et al., 2004). Before being able to engage in argumentation, students must develop a sense of how to critically analyze evidence and its relation to claims. Arguments are essential to the process of justifying the validity of any explanation as there are often multiple theories for any given phenomenon (Osborne & Patterson, 2011). Within the field of science, arguments take place at multiple levels: within the mind of individual scientists and across research groups where alternative directions are considered (Driver, et al., 2000).
Analysis on the quality of arguments diverges in two directions: first student ability to assess arguments and the methods students use to assess arguments, and second, teacher assessment of arguments. This differs from the structure of arguments because analysis of the structure of arguments focuses on the components that make up an argument while research into quality of arguments explores the cohesion of an argument as a whole—how the components in an argument work together to create a product. These analyses include adaptations of the Toulmin model specifically for science education research regarding rebuttals as counterarguments and qualifiers (i.e., presumably, always, almost certainly) to put forth high quality arguments (Lee et al., 2014). Criteria for evaluating the quality of arguments can be empirical, theoretical, and analytical, probing for the cohesion of an argument as a whole in terms of both process and content (Sampson & Blanchard, 2012; Simon et al., 2006). These studies find that students are unable to evaluate the credibility of evidence and that this occurs throughout all levels of education as a result of a lack of knowledge of the criteria with which to use to judge credibility (Nicolaidou et al., 2011; Driver, et al., 2000).

2.4.3 Argument as process.

Studies focusing on argumentation in the social context of the classroom use other frameworks for analysis which are based on Vygotskian principles of science teaching, taking into consideration the social construction of knowledge: that argumentation is a negotiated social act within a specific community (Scott et al., 2006; Driver et al., 2000). Studies on the social component of argumentation focus on the process of argumentation rather than the product, which is an argument. These studies analyze discourse, teacher
questioning, and student collaborative arguments (Chin, 2007; Chin & Osborne, 2010; Christodoulou & Osborne, 2014; Berland & Hammer, 2012).

Analysis of discourse in argumentation looks at the nature of discussions that take place in the classroom. This includes argumentative processes such as student contributions, spontaneity of students’ participation, and statement, defense, evaluation, and revision of claims through science ‘talk’ (Berland & McNeill, 2010). These studies view argumentation as a verbal, social, rational activity where co-construction of knowledge is either dialogic or rhetorical (Lee et al., 2014). Understanding the co-construction of knowledge requires consideration about students’ common ground, their prior beliefs and assumptions, and an analysis of the ways in which a space is negotiated for constructive talk that draws on what students encounter in their everyday lives as a resource for learning science (Clark & Schaefer, 1989; McNeill 2011). The attention is foremost on creating a classroom culture that provides a stage for reflective discourse.

Research also looks at why some students are more successful at providing high quality arguments. Collaborative argumentation occurs in a space where “learners are expected to share ideas, question assumptions and restructure their existing knowledge schemata based on the interactions in their group” (Evagorou & Osborne, 2013, p. 212). These studies are guided by works in science classroom discourse, argumentation discourse, and theories about social construction of knowledge in relation to scientific practices. Findings are that students are generally unable to select appropriate data as evidence when they are asked to participate in argumentative discussion (Berland & McNeill, 2010). When students participate in collaborative argumentation, they frequently
are unaware of the reasoning or purpose behind this constructive act and do not relate it to scientific practice.

2.5 Argumentation and the Nature of Science

Recent studies in arguments explore the nature of science and its application to arguments. Toulmin (1984) put forth an adapted model for scientific arguments. In his framework, there are certain cause and effect relationships taken to be routine or expected. Anomalies in nature are what science seeks to explain. Explanations are used to put to rest scientific anomalies. There are four methods through which explanations can be put forth: explanation by type, material composition, history, and goal. These explanations vary in what components of argument (grounds, claims, rebuttals, warrants) are necessary and sufficient to account for the anomaly. However, explanations alone are not enough to satisfy the needs and criteria for acceptable scientific knowledge. A justification must also be included. This justification responds to the question “What does the solving of this particular problem do for science,” specifically within the context of what science seeks to accomplish (Toulmin, 1984, p. 328). While this framework considers explanations as resolutions to scientific anomalies, scientific arguments are used either to challenge the application of a theory or to put forth alternative or refined theories in an iterative process similar to science practices.

Toulmin goes on to define what counts for each specific component of an argument. Claims are factual statements that are drawn from the other components of a regular scientific argument. Grounds of an argument can vary from data centered evidence to deeper considerations of the history and context of the claim to be drawn. Warrants justify
the connection from grounds to claims, consisting of something as concrete as a mathematical equation describing a known result of phenomena to abstract laws of nature. Backings provide specific justification for the warrant including systematic, empirical support aligned with warrants. Models and rebuttals, similar to the field-invariant model, are circumstances under which the argument does not hold true.

An additional type of scientific argument included in this framework is a critical scientific argument. Rather than focusing on objects and processes, critical arguments focus on the theories behind them. Critical scientific arguments have particular standards for the components of an argument, much like the regular scientific arguments stated previously. They are pragmatic by nature and include justifications on the line of the enterprise of science. In critical scientific arguments, warrants show how the grounds (theories) fit into the general purpose of the nature of theories. These arguments focus more on the processes accompanying scientific arguments and whether these processes and products align with scientific practices.

2.6 Arguments as a Social Practice

In order to promote socially mediated construction of knowledge, argumentation has been regarded as a method to help enhance student critical thinking and conceptual understanding through discursive process. This understanding of argumentation comes from two perspectives: rhetoric and logic (Wenzel, 1990). Rhetorical theorists conceptualize argumentation as a process of using language and other representations to influence and persuade people. This view is very important in politics, business, and law, but often less so in the science classroom because of the positivist perspectives that tend to
dominate (Toulmin, 1984). Logicians examine the logical structure of arguments. The logical soundness of arguments is often evaluated using the Toulmin (1958) model. Studies in student argumentation have cited the Toulmin model as a basis for rubrics to analyze written products of argumentation (Erduran, Simon, & Osborne, 2004, Jimenez-Aleixandre, Rodriguez, & Duschl, 2000, McNeill & Krajcik, 2008). Linguists study the conversational interactions among people who hold opposing views (Van Eemeren & Grootendorst, 2004; D. Kuhn & Udell, 2003). It is a general consensus that argumentative conversations should be goal-directed: the goal of having an argumentative conversation is to resolve the difference of opinion rather than win the debate. Therefore, the participants of the conversation must make appropriate discourse moves to exchange, evaluate, and revise viewpoints (Eemeren & Grootendorst, 1987, 2004; D. Kuhn, & Udell, 2003). This notion of goal-directed argumentative conversations is consistent with the nature of science. In the history of science, scientists argued about competing theories, facts, laws, and inferences. As a result of their goal-directed conversations, new theories were generated to replace the accepted ones in order to keep the commensurability of the theories in the field (Kuhn, 1962). The idea of goal-directed argumentative conversations is important because it opens the discussion to the distinction between argumentation and explanation (Berland & McNeill, 2012; Osborne & Patterson, 2011).

Researchers have begun to conduct substantial work to investigate students’ argumentation practice. Past studies have explored the connection between conceptual understanding and the understanding of the logical structure of arguments in online environments (e.g. Sampson & Clarke, 2008), the understanding of argumentation in classroom settings (e.g. Osborne, Erduran, & Simon, 2004), student argumentation (e.g.
Sampson & Clarke, 2008; Jimenez-Aleixandre et al., 2000; Osborne et al., 2004), and teacher pedagogical strategies to help facilitate argumentation in the classroom (e.g. Simon, Erduran, & Osborne, 2006; McNeill & Krajcik, 2008). Studies have shown that students had difficulties in providing relevant evidence, using evidence to justify claims, and rebutting counterarguments (e.g. Berland & Reiser, 2009; McNeill, 2011; Osborne, Erduran, & Simon, 2004).

2.7 The Role of Argumentation in Science Classrooms

Previously, arguments have been included in the curriculum as forms of written explanations. Berland and McNeill (2012) contend that argumentation and explanation have many similarities and, by exploring these similarities, argumentation can be easily integrated into existing curriculum. The parts of explanations and arguments serve similar functions in educational settings. For example, creating an explanation for a conclusion based on evidence is similar to creating justification for a claim based on evidence in argumentation. The goal for teaching students to make explanations based on evidence is to help students create a knowledge base, which prepares them to be able to explain scientific phenomena (Berland & McNeill, 2012).

However, other scholars have pointed out that argumentation varies from explanation in that argumentation is about justifying conclusions that are equivocal and uncertain (Osborne & Patterson, 2011). Since the difference of opinion on uncertain issues can only be resolved through exchanging, evaluating, and revising different viewpoints, the notion of goal-directed conversation is crucial to argumentation but not central to explanation. The push towards inquiry science has endorsed the role of explanation in
classrooms. Explanation has been taught as a technique to provide summaries that supplement the content being taught in class.

Given the synergistic relationship between argumentation and explanation, the present inclusion of argumentation in the curriculum can be viewed as a matter of adapting aspects of explanation to help include the added elements of argumentation and more direct goals associated with teaching argumentation.

Refining the process of creating an explanation, is a shift towards the learning goals of argumentation (Abi-El-Mona & Abd-El-Khalick, 2006; Cavagnero, 2010). The NRC Framework (2011) justifies the need to include argumentation as a skill to teach in the science classroom because it serves its purpose in preparing students for careers in science as well as preparing students to become well-informed citizens. In preparing students for careers in science, being able to identify weaknesses in claims and learning to give the best possible explanation to resolve questions, techniques for data analysis, and interpretation of data is a practical skill (NRC, 2011). In preparing students to become well-informed citizens, the ability to identify weaknesses and limitations in claims help students sift through “bad science” (NRC, 2011, p. 71) and be critical of science-related reports found in their day-to-day experiences.

Argumentation can be used as a tool beyond simply looking at the ways in which evidence connects to claims. By having students explicitly state the warrant that connects the evidence to the claim, students have to address scientific principles (backings) that help to support their reasoning. Through this process, students gain a stronger understanding of the foundational principles that are taught as content in the course. The way to address the weakness in students’ understanding of argumentation is to teach it as part of the
curriculum (Carelas, Pappas, Kane, Arsenault, Hankes, & Cowan, 2008; Osborne, Erduran, & Simon, 2004).

The purpose of teaching argumentation as part of the curriculum is to give students an understanding of the importance in the connections between the parts of creating an argument (Cavagnetto, 2010; Jimenez-Aleixandre, Rodriguez, & Duschl, 2000). New methods for using argumentation in different scientific concepts are also being developed because curriculum also has an impact on the way that argumentation should be presented to the class (Varelas et al., 2008). Different content topics require different methods for using argumentation (Varelas et al., 2008).

Engaging in arguing-to-learn depends on students’ abilities in both argumentative discourse and content-based arguments. Teachers should provide supports to encourage both dimensions during argumentation lessons. For example, by using sentence stems and visualization tools, such as the diagram that helps students visualize the structure of scientific arguments, teachers can engage students in argumentative conversations that draw on scientific content knowledge and effectively exchange and evaluate different views. Through these conversations, students collectively construct scientific mechanisms and develop conceptual understanding of scientific concepts and principles.

Kuhn and Reiser (2006) developed a rubric for assessing levels of argumentative skill. The finding resulted in the knowledge that teachers must develop curriculum explicitly to teach argumentation. This requires teaching a combination of the process of argumentation and the product of argumentation. McNeill and Krajcik (2008) explored how teachers’ different uses of the explanation framework in their classrooms influenced student learning. The framework consisted of four instructional practices: defining
scientific explanation, making the rationale of scientific explanation explicit, modeling scientific explanation, and connecting scientific explanation to everyday explanation. Students had the most trouble with reasoning, which allowed for the most room for growth.

2.8 Evaluating Arguments

Evaluating arguments has a variety of classifications and analyses in argumentation research studies. Some use Toulmin’s structural model, by looking at whether arguments contain specific components such as claims supported by data with warrants that connect the two and include backings, and qualifiers. Others analyze arguments in terms of the issues they address such as definition (what is an argument?), causal (how did it get this way?), evaluation (is it good or bad?), and proposal (what should be done about it?).

Evaluative analyses on argumentation intervention are divided into three domains: nature of the argument intervention, emphasis of the argument intervention, and aspects of science included in the argument intervention in order to observe the ways research is distilled into practice (Cavagnetto, 2010). Studying the nature of the argument intervention places focus on a distinction in the phase of instruction in which argument was integrated and the extent to which argument was used for gaining new understanding.

2.8.1 Discourse in argumentation.

Discourse in argumentation has multiple points of focus. The first analyzes student participation in argumentation. These studies explore the types of talk and strengths and weaknesses found in student argumentation. The second point of focus is on teacher facilitation for argumentation which explores how teachers encourage argumentation in the
classroom. The third classifies argumentation as either an individual activity or a social activity. Sampson and Blanchard (2012) define argumentation as “a knowledge building and validating practice in which individuals attempt to establish or validate a conclusion, explanation, conjecture, or other claim on the basis of reasons” (p. 1123). The various points of focus in argumentation are reflected within this definition. Argumentation is a negotiation of knowledge that includes putting forth concepts and content and critically analyzing explanations or claims. In this sense, discourse in argumentation requires student participation and the teacher scaffold. More connections need to be made on how teachers facilitate and evaluate written and oral discourse.

2.8.2 Student argumentation

Research in student argumentation investigates analysis of student argumentation ability and the written discourse of an argument, developing models created evaluation (e.g. Evagorou & Osborne, 2013) and analyses of classroom discourses (e.g. Jimenez-Aleixandre, Rodriguez, & Duschl, 2000). Additionally, studies explored connections between student argumentation and conceptual understanding, showing argumentation discourse has a significant impact on student concepts in the nature of science, conceptual understanding, and participation in school.

Jimenez-Aleixandre, Rodriguez, and Duschl (2000) investigated argumentation as a social activity, analyzing the classroom as a unit and the discourse involved between students within the classroom. Building upon concepts about procedural displays in school, Jimenez-Aleixandre, Rodriguez, and Duschl (2000) reviewed audiotapes of student dialogue and broke them up into units of analysis including scientific culture and school
culture; “talking science” and “doing science,” argumentative operations in the discourse and epistemic operations relevant to the development of scientific knowledge. They found that student dialogue tends to be dominated by school culture based on students’ need to “do school” over “doing science.” Within the classroom, there are issues arising in terms of co-constructed arguments and unbalanced participation. For example, some students contributed more to the dialogue than others and students had varying degrees of argumentation ability. Students tend to use analogies as support in discussions, incorrect backings, and supports through anthropocentrism when developing arguments and explanations (Jimenez-Aleixandre, Rodriguez, & Duschl, 2000).

In previous studies, researchers have developed different ordinal scales to evaluate students’ arguments about scientific and socio-scientific issues (e.g. Berland & McNeill, 2010; Clark & Sampson, 2007a; Erduran et al., 2004). Usually, arguments at the lowest level only contain a claim, whereas arguments at a higher level contain evidence and/or warrant. Arguments at the highest level contain all argument elements, including claims, evidence, warrants, and rebuttals. The levels capture student development in understanding the logical structure of arguments.

2.9 Reasoning and Explanations

Past works on reasoning begin by studying rhetoric, logic, and reasoning, similar to logos in argumentation schemes. Rhetoric is how logic is packaged and logic is the study of reasoning (Brinkley, 1995). Through the process of reasoning an individual comes to a “reckoning,” which is described in current studies as a theory. Current studies define reasoning, arguments, and theories as related entities. In these studies, the demarcation
between theories and reasoning is more difficult to draw. Most commonly, studies regard reasoning as a process through which arguments are created, implying that argumentation and scientific reasoning are completely intertwined notions, able to be used interchangeably (e.g. Jimenez-Aleixandre et al., 2000; Nicolaidou et al., 2011; Cavagnetoo, 2010; Osborne & Patterson, 2011; Berland & McNeill, 2012; Driver et al., 2000).

A framework from Lee et al. (2014) describes scientific reasoning as a component of scientific argumentation and critical reasoning as an analysis of strength of an argument, contending that the critical reasoning side has been widely neglected. Their study focuses on reasoning as an individual’s ability to coordinate between theory and a critical assessment and thinking related to the strength of an argument (Lee et al., 2014). However, there are studies that consider reasoning to be a theoretical process within itself and argumentation to be a form of reasoning. Some studies find that “argumentation is held to be a reasoning strategy and, thus, also comes under the reasoning domains of informal logic and critical thinking,” implying that argument is a component of reasoning, rather than reasoning as a component of argumentation (Jimenez et al., 2000, P. 760).

2.10 Frameworks for Reasoning

Reasoning is described as a coordination of claims and evidence, an implicit process (Sandoval & Millwood, 2005). Rather than considering reasoning as a general theory under which processes of argumentation fall, this study looks at arguments as a general process encompassing many components, one of which is reasoning. Argumentation includes an internal reasoning process influenced by epistemic beliefs and
coordination of claims and evidence. With this consideration, reasoning takes place as a piece of an argument, strengthening its persuasive aspect and putting forth an explanation that justifies the argument. Therefore, explanations, arguments, and reasoning are linked such that implicit reasoning takes place as a part of argumentation, which is a process that includes implicit and explicit reasoning, resulting in an explanation. Figure 3 provides the summary from the framework which provided the basis of exploration in reasoning for this study.

<table>
<thead>
<tr>
<th>Type of representation of informal reasoning</th>
<th>Study</th>
<th>Main results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision-making modes</td>
<td>Sadler &amp; Zeidler (2005)</td>
<td>Three patterns of informal reasoning: rationalistic, emotive, and intuitive</td>
</tr>
<tr>
<td></td>
<td>Patronis et al. (1999)</td>
<td>Four modes of arguments: social, ecological, economic, and practical</td>
</tr>
</tbody>
</table>

*Figure 3. Summary of informal reasoning (from Wu and Tsai, 2007, p. 1167)*

Some studies differ by classifying reasoning into two distinct categories: informal and formal reasoning. In a Wu and Tsai (2007) study, reasoning is described as the process of constructing and evaluating arguments. The framework for this study is built upon multiple theories for rational human thought including the Wu and Tsai (2007) framework.

Generally, reasoning is described to be either inductive (bottom up) or deductive (top down). In bottom up, or inductive reasoning, individuals detect patterns and regularities from observations and a theory for explaining the phenomenon is developed.
By comparison, in top down, or deductive reasoning, theories are first developed through prior knowledge and used to design a testable hypothesis. Observations testing these hypotheses are documented leading to acceptance or refutation of the theory.

Using Kuhn and Pearsall’s (2000) framework on scientific thinking, theories contain beliefs and/or knowledge claims and either one factor as cause or a system of interconnected factors as explanations. Theories are developed through top down and bottom up reasoning processes. These two combined processes help an individual negotiate through their prior knowledge and beliefs to develop the most plausible explanation or claim to explain a phenomenon.

2.11 Community College Students

The population of students explored in this research study are students in community college. Community colleges were originally designed to help make higher education more attainable for people with lower economic statuses, providing training similar to the vocational schools of today (Cohen, 2003). The majority of students attending community colleges were those who were not served by higher education institutes (Hill, 2016). Today, community colleges have expanded their target population to include students preparing for terminal degrees in higher education institutes. The demographics the community college used for this study had an equal percentage of students aged 17-21 as students above the age of 45. Students in community colleges encompass a wide range of variation in terms of educational background, purpose for attending college, age, and personal experiences (Kopko & Crosta, 2016). This under-researched population consisting of students with a wide range of experiences would provide robust data that will
help contribute to the greater knowledge base for educational research by providing more insight on how personal experiences impact learning processes.

2.12 Conclusion

This literature review has discussed the differences between classroom science and the nature of science practice, explanations, argumentation, and reasoning. There is a gap between science practice and classroom science. Current science education research seeks methods to bridge the gap. Argumentation and reasoning have been put forth as possible building blocks to close the gap. While many studies analyze social practices of argumentation, explanation, and reasoning, very little exploration has been conducted on students’ internal processes. While this research study does not seek to provide an absolute answer or understanding about students’ internal reasoning and argument building process, it does seek to provide some insight. The first two assessment tasks were used to investigate the connections between reasoning abilities and understanding of arguments. The open response and interview task elicited students’ argument development process and methods for justifying claims.
Chapter 3: Methods and Procedures

The purpose of this study was to explore connections between reasoning and arguments and to gain insight into students’ reasoning and justification processes. To explore reasoning and argumentation, this study used a combination of surveys and interviews. Each participant completed the Lawson Classroom Test of Scientific Reasoning (LCTSR), the Test of Scientific Argumentation (TSA) and an open response questionnaire (OR) which uses 3 selected questions from the Conceptual Inventory of Natural Selection (CINS). The 3 selected CINS questions are posed to students on the questionnaire as open-ended Participants of this study were given two assessments and an open response questionnaire. A small subgroup of students was interviewed to elicit more information about the processes leading to their responses to the open response questionnaire. Statistical analyses were conducted on sub-scales of the LCTSR and TSA to explore connections between students’ reasoning skills and argumentation. A coding scheme for assessing the open response questionnaire was developed using frameworks from existing literature on argumentation. Responses to the questionnaire were coded based on those frameworks and emergent patterns were also documented. Interviews were analyzed and coded to find patterns in students’ process of analyzing counterarguments and justifying arguments.
The research questions guiding this study are:

1) How do community college students perform in reasoning and understanding the parts of an argument and how do these skills correlate?

2) What kinds of claims do community college students use when developing arguments?

3) How do community college students justify and/or defend their claims when presented with a counterargument?

This study included 21 participants taking classes at an Associate’s Degree granting community college, a population that encompasses a wide range of age, ethnicity, and career goals. However, despite the wide variation across the population, very little research has been conducted focusing on this group. This chapter discusses the participants, the data collection procedures, and the coding scheme for both the open response questionnaire and corresponding interviews.

3.1 Participants

Participants in the study were all students attending a community college that has an open enrollment policy. In open enrollment, the only admissions requirement is a high school diploma or equivalent (Sinclair Community College, 2017). This population includes twice as many part-time enrolled students as full-time students, with 47% of students between the ages of 18-24 and 47% of students between the ages of 25-64 (Sinclair Community College, 2017). Despite community colleges being known as an alternate pathway to a baccalaureate, the retention rate and degree completion across a 6-year period
is lower for university transfer students than for students in the 4-year university cohorts (Ohio Articulation and Transfer Network, 2014).

The sample consisted of 21 students from Physics, Chemistry, and Math courses. Each class consisted of approximately 15-18 students. Out of 48 total students, only 21 fully completed all three assessments, a 44% return rate. The Physics classes were College Physics and Physics for Technology. College Physics serves as an introductory course and general education requirement. Physics for Technology is a required course for licensure in programs such as physical therapy, radiology, dietician, and dental hygiene. The Chemistry course was also an introductory, general education requirement for science majors (associate in science) and students applying to transfer to a 4-year institution for a Bachelor of Sciences in Biology or Chemistry. The math class was an intermediate algebra class, which is a first-year math class that can be taken by both science and non-science majors. The sample included a variety of majors (see Table 1)

<table>
<thead>
<tr>
<th>Major</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemistry</td>
<td>2</td>
</tr>
<tr>
<td>Physical Therapy Assistant</td>
<td>4</td>
</tr>
<tr>
<td>Business Administration</td>
<td>1</td>
</tr>
<tr>
<td>Unmanned Aerial Systems</td>
<td>1</td>
</tr>
<tr>
<td>Engineering Transfer</td>
<td>2</td>
</tr>
</tbody>
</table>

Continued
Table 1 continued

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mechanical Engineering</td>
<td>1</td>
</tr>
<tr>
<td>Biology</td>
<td>3</td>
</tr>
<tr>
<td>Advanced Technical Intelligence</td>
<td>1</td>
</tr>
<tr>
<td>Professional Pilot and Airway Science</td>
<td>1</td>
</tr>
<tr>
<td>Clinical Laboratory Science</td>
<td>1</td>
</tr>
<tr>
<td>Dermatology</td>
<td>1</td>
</tr>
<tr>
<td>Dietician</td>
<td>1</td>
</tr>
<tr>
<td>Unspecified</td>
<td>2</td>
</tr>
</tbody>
</table>

The following tables summarize demographics of the participants in this study, which included 11 females and 10 males, ages ranging from 18 to 57 and number of years in college ranging from 1 to 8 years. Table 2 presents participants’ ethnicities, Table 3 contains the number of college level science classes previously taken by participation, and Table 4 gives a breakdown of the number of college level math classes previously taken by participants. The majority of participants only had 1-5 college math classes and 1-5 college science classes. These math classes included development/remediation courses such as basic arithmetic, basic mathematics, and introduction to algebra. Past science classes included anatomy and physiology, physics, introduction to chemistry, and biology.
Table 2. Participants' ethnicity

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian Indian</td>
<td>1</td>
</tr>
<tr>
<td>Black or African American</td>
<td>3</td>
</tr>
<tr>
<td>Caucasian or White</td>
<td>12</td>
</tr>
<tr>
<td>Latino or Hispanic</td>
<td>1</td>
</tr>
<tr>
<td>Table 2 continued</td>
<td></td>
</tr>
<tr>
<td>More than one race</td>
<td>2</td>
</tr>
<tr>
<td>Unspecified</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
</tr>
</tbody>
</table>

Table 3. Number of college level science classes previously taken

<table>
<thead>
<tr>
<th>Number of Science Classes</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>16</td>
</tr>
<tr>
<td>6-10</td>
<td>4</td>
</tr>
<tr>
<td>16 or more</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
</tr>
</tbody>
</table>

Table 4. Number of college level math classes previously taken

<table>
<thead>
<tr>
<th>Number of Math Classes</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>17</td>
</tr>
<tr>
<td>6-10</td>
<td>2</td>
</tr>
<tr>
<td>11-15</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
</tr>
</tbody>
</table>

3.2 Data Collection and Assessment Instruments

This study had two phases. In the first phase, students were given written assessments. These written assessments were validated instruments, commonly used in education research. In the second phase, students were interviewed for a more in-depth understanding of their responses to the open response questionnaire.
3.3 Written Assessments

The three written assessments given to participants were: The Test of Scientific Argumentation (TSA), the Lawson Classroom Test of Scientific Reasoning (LCTSR), and an Open Response Questionnaire (ORP) developed from the Conceptual Inventory on Natural Selection (CINS). The TSA and LCTSR are validated instruments (Lawson, 1978; Frey, Ellis, Bulgren, Hare & Ault, 2015). Validity and reliability of the tests will be presented in subsequent sections. The ORQ includes three question stems from CINS as open response items rather than multiple-choice items. Students were given time during class to complete all written assessments. These assessments took about 40-50 minutes to complete.

3.3.1 Test of scientific argumentation.

The Test of Scientific Argumentation (TSA) was developed to provide researchers, teachers, and program evaluators a tool to analyze participants’ understanding of the parts of an argument (Frey, Ellis, Bulgren, Hare & Ault, 2015). This assessment tool was selected for this study because it is a recently developed assessment that specifically targets parts of arguments in domain general contexts. The constructs measured by this assessment are distinguishing between claim, fact, data, and opinion, identifying qualifiers, labeling statements as claims or non-claims, ranking quality of reasoning, and labeling grounds as authority, logic or theory, and looking at rebuttals and counterarguments. While this study does not look into students’ abilities on all of these different areas, the TSA includes tests of three main structures of interest pertaining to this study: claims, grounds, and counterarguments. In this assessment, claims are defined as statements including scientific
observations leading to a conclusion about the natural world. Statements are intended to persuade another person (Frey et al., 2015). Counterarguments are alternate claims that are based on reasoning and evidence. Logic, authority, and theory are reasons for believing a claim. A theory is a set of organized statements explaining phenomena while authority uses an appeal to trusted sources of information and logic applies to sets of rules applied to for making conclusions (Frey et al., 2015). Arguments include any combination of claims in addition to logic, authority, and/or theory which support the claim. For this study, these structures of arguments are similarly defined--logic, theory, and authority make the grounds of an argument, which are reasons why people believe statements. Understanding students’ abilities in identifying, distinguishing, and labeling these structures will inform on the later part of the research study. Scores on the TSA show whether students are able to identify these structures of an argument. For example, if students score low on the section asking them to identify counterclaims, then we could anticipate that students would have weaknesses identifying counterclaims in others’ arguments or have trouble defending against counterclaims.

Questions on the TSA are scored on a 36-point scale, with one point per item answered correctly. These items are sorted into six general categories: labeling statements as claim, fact, data, or opinion (corresponding to items 1-6); finding qualifiers in statements (corresponding with items 7-12), deciding if a statement is a claim or not a claim (corresponding with items 13-18), determining whether the grounds of an argument are supported by authority, logic, or theory (corresponding with items 19-24), labeling statements as rebuttals or counterarguments (corresponding with items 25-30), and recognizing the quality of reasoning (corresponding with items 31-36).
For test design, an initial pool of items was developed by content experts and then rated independently by members of the research team in terms of how well they provided examples of the intended construct (Frey et al., 2015). A subset of items that had the highest ratings were pilot tested to middle school science students as a paper and pencil test. Concurrent validity was measured by correlating scores on this test with the Cornell Thinking Test and the TSA was deemed valid with a moderate correlation \( r=.59, n=54, p<.001 \) across the tests (Frey et al., 2015). The overall difficulty level of this test is .68, indicating moderate difficulty and Cronbach’s alpha of .82 indicating good reliability (Frey et al., 2015). The Test of Scientific Argumentation can be found in Appendix A.

There are two major factors that may impact how students develop arguments. The first is their understanding of the structures of arguments and the second is their reasoning abilities. The TSA measures the former and the LCTSR measures the latter.

### 3.3.2 Lawson classroom test of scientific reasoning.

This study defines domain general reasoning as the justification process used to defend arguments. In science thinking, these “formal operations include those reasoning processes that guide the search for and evaluation of evidence to support or reject hypothetical causal propositions” (Lawson, 1978, p.12). In this definition, there are a few ways of thinking that dominate reasoning: 1) hypothetical-deductive; 2) correlational; 3) identification and control of variables, 4) probabilistic thinking; and 5) proportional thinking. The Lawson Classroom Test of Scientific Reasoning tests students’ thinking in these categories (Lawson, 1978).
The Lawson Classroom Test of Scientific Reasoning (LCTSR) contains multiple-choice questions concerned with key skills that are linked to scientific reasoning (Lawson, 1978). The LCTSR is conventionally used in education research as a tool for understanding students’ reasoning abilities and has been shown to be valid and reliable (Lawson, 1993). Items in the LCTSR are scored on a 24-point scale with one point per correct item. This test assesses the following types of reasoning: conservation of weight (items 1 and 2), conservation of displaced volume (items 3 and 4), proportional thinking (items 5-8), identification and control of variables (items 8-14), probabilistic thinking (items 9-18), correlational thinking (items 19 and 20), and hypothetical-deductive thinking (items 21-24).

In this study, the items associated with understanding conservation of weight and conservation of displaced volume are not included because they are more related to specific content in science. The remaining items related to scientific thinking (proportional thinking, identification and control of variables, probabilistic thinking, correlational thinking, and hypothetical-deductive thinking) are analyzed. In proportional thinking, comparisons between two entities are drawn in multiplicative terms, including being able to apply ratios from one situation to another. Identification and control of variables tests students’ ability to identify independent and dependent variables when testing hypotheses. Correlational thinking is the reasoning process involved in determining the strength of relationships between variables. Hypothetical-deductive thinking incorporates developing and testing possible solutions for problems.

The LCTSR is an appropriate measure to explore student thinking within the frame of this study because justification processes used to defend scientific arguments are
influenced by ways of science thinking and the LCTSR measures these constructs. The interview portion of this study provides qualitative data that can be used to understand student performance on the LCTSR. For example, if students score high on hypothetical-deductive thinking, when asked for justification for their arguments, students may do so by focusing on the nature of their own claim or attacking the counterclaim. Students scoring well on correlational thinking are better at identifying patterns in data and may focus on evidence when encountering counterarguments.

For test validation, the LCTSR was given to a large sample of eighth, ninth, and tenth grade students (Lawson, 1978). A subsample of this group was randomly selected for interview on four Piagetian tasks including conservation of weight, displaced volume, bending rods, and balance beam. These interview tasks included all reasoning processes targeted on the LCSTSR. The LCTSR was deemed valid because there was a strong positive correlation in scores on the LCTSR and the interview task and was considered reliable with a Kuder-Richardson (KR-20) score of .78 (Lawson, 1978). The LCTSR is a widely-accepted assessment for exploring reasoning. The Lawson Classroom Test of Scientific Argumentation can be found in Appendix B.

### 3.3.5 Open response questionnaire.

The Open Response Questionnaire (ORQ) has four total questions. The first question is not framed within any specific science content and asks students to identify connections between three variables (fishing rod length, fish hook thickness, and location) in a fishing scenario. In the prompt for this question, students are given a data table that shows the number of fish caught by three different individuals with different combinations
fishing rod length, fish hook thickness, and location. The remaining three questions are framed in natural selection. These questions were drawn from the Conceptual Inventory on Natural Selection (CINS) and focus on differential survival, origin of variation, and origin of species. While the CINS is a multiple-choice assessment, the open response questionnaire in this study gave the question stems as open-ended to elicit responses unrestricted by the four choices given in the CINS. These open-ended questions allow for more flexibility in responses. The following sections discuss each question on the open response questionnaire in more depth. This questionnaire can be found in Appendix C.

3.3.4.1 Open response question 1.

Question 1 contains the following prompt (see Figure 4).

1. Tom, Jerry, and Dan are good friends and go fishing together most weekends. They often use the same type of fishing tools and have similar skills in fishing (that is, they often each catch a similar number of fish every time). On their last fishing trip, they had selections of different fishing rods and fishhooks, and they each picked a different location to fish (see the conditions given in the table below). They fished for a total of two hours and the number of fish they each caught during this period is shown below.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Tom</th>
<th>Jerry</th>
<th>Dan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fishing rods</td>
<td>Long</td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td>Fishhooks</td>
<td>Thick</td>
<td>Thin</td>
<td>Thin</td>
</tr>
<tr>
<td>Locations</td>
<td>Point A</td>
<td>Point A</td>
<td>Point B</td>
</tr>
<tr>
<td>Number of fish caught during the two-hour</td>
<td>15</td>
<td>15</td>
<td>8</td>
</tr>
</tbody>
</table>

Figure 4. Open response questionnaire, question 1

Because this question was not framed within a science content, students’ reasoning would not be inhibited by content knowledge. Data are given in the table so students have all the information they need to respond to the question. They could use data from the table,
personal experiences, or other knowledge as reasoning for their responses. In this question, students were asked: a) What can you say about the link between each of the three conditions and the number of fish caught and b) why do you say that? The purpose of part A was to ask students for a claim about the data and part b asked for the evidence supporting their claim.

### 3.3.4.2 Open response question 2.

The second question contains the following prompt (see Figure 5).

“*The Canary Islands are seven islands just west of the African continent. The islands gradually became colonized with life: plants, lizards, birds, etc. Three different species of lizards found on the islands are similar to one species found on the African continent (Thorpe & Brown, 1989). Because of this, scientists assume that the lizards traveled from Africa to the Canary Islands by floating on tree trunks washed out to sea*" (Anderson, Fisher, & Norman, 2002, p. 974).

Fitness is a term often used by biologists to explain the evolutionary success of certain organisms. Below are descriptions of four fictional female lizards.

<table>
<thead>
<tr>
<th></th>
<th>Lizard A</th>
<th>Lizard B</th>
<th>Lizard C</th>
<th>Lizard D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body length</td>
<td>20 cm</td>
<td>12 cm</td>
<td>10 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>Offspring</td>
<td>19</td>
<td>28</td>
<td>22</td>
<td>26</td>
</tr>
<tr>
<td>surviving to</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adulthood</td>
<td>4 years</td>
<td>5 years</td>
<td>4 years</td>
<td>6 years</td>
</tr>
<tr>
<td>Age at death</td>
<td>4 years</td>
<td>5 years</td>
<td>4 years</td>
<td>6 years</td>
</tr>
<tr>
<td>Comments</td>
<td>Lizard A is very healthy, strong, and clever</td>
<td>Lizard B has mated with many lizards</td>
<td>Lizard C is dark colored and very quick</td>
<td>Lizard D has the largest territory of all the lizards</td>
</tr>
</tbody>
</table>

*Figure 5. Open response questionnaire, question 2*

This question contains a question stem from the Conceptual Inventory on Natural Selection. The concept tested is differential survival. Hereditary makeup of individuals influences survival and individuals with qualities that help them survive in their
environment are likely to leave more offspring than individuals that are less fit (Anderson et al., 2002). For this two-part question, students were asked a) to identify which lizard biologists would deem most fit and b) to explain how and why they came to that conclusion. This question was selected from the CINS because its format is similar to the previous question, giving students a table that can be used as evidence in their arguments. If students did not have content knowledge support on differential survival, they could still construct responses based on their personal experience(s) and/or knowledge as well as the data given in the table. Students were asked to explain both how and why they came to their conclusion, which prompts for both evidence and warrants to explain their claims.

### 3.3.4.3 Open response questions 3 and 4.

Content for question 3 involved the origin of variation. Isolated populations, over time, may change so much that it becomes a new species (Anderson et al., 2002). Question 4 involved the origin of species. Random mutations and sexual reproduction can produce variation and mutations can be harmful, have no significant effects, or have benefits in some environments (Anderson et al., 2002). To be adapted for this study, both were written as questions containing two parts. Figure 6 shows the prompt for question 3. Figure 7 is the prompt for question 4.

3. According to the theory of natural selection, from where did the variations in body size in the three species of lizards most likely come? Please explain how and why you came to that conclusion.

*Figure 6. Open response questionnaire, question 3*
4. What could cause one species to change into three species over time? Please explain how and why you came to that conclusion.

Figure 7. Open response questionnaire, question 4

In these open-ended questions, students’ responses contained classroom knowledge, personal knowledge and experiences, or a combination of both classroom and personal knowledge.

3.4 Interview Protocol

In the interview portion of the study, students discussed their responses to questions 1 and 2 from the open response questionnaire. Students were given their completed open response questionnaires for reference during the interviews, which were conducted within 2 weeks of students completing their written assessments. Only questions 1 and 2 were selected for the interview questions because they provided the most opportunities for probing through questioning. Students did not need a strong understanding of evolutionary processes and natural selection to respond to these prompts.

The interviews followed a semi-structured protocol. In a semi-structured interview protocol, the interviewer has some general questions shaping the interview. However, the flow and direction of the interview is guided by the content of the responses. Thus, semi-structured interviews have a general interview protocol but the interviews can vary across participants. The questions included in every interview were prompts for various parts of arguments including claim, evidence, warrant, process, and other factors (see Figure 8). The prompt for claim asks students for a statement or conclusion drawn from the data in
tables. If students did not provide evidence with the prompt for claim, they were asked for evidence. This allowed for exploration about what students counted as evidence and how much they felt was both necessary and sufficient to support their claims. The prompt for warrant was used to see how students would connect their claims and evidence. In the framework for this study, warrants are viewed as reasoning used to justify claims. An example of this type of warrant is setting grounds and contexts for an argument. The prompt for process was included to explore whether students have metacognition of how they view structures of the argument or whether they are cognizant of their process of manipulating the structures and putting them together to create a cohesive argument. Finally, the prompt for other factors looks at whether there were any outside variables that influenced their thought process when developing their claims. This could include personal experiences or experiences outside the classroom.

**Prompt for claim:** What did you say were the connections between fishing rods, locations, fish hooks, and number of fish caught?
**Prompt for evidence:** Why do you say that?
**Prompt for warrant:** How do you know?
**Prompt for process:** When approaching this question, what initially caught your attention?
**Alternative prompt for process:** Tell me a little bit about your thinking as you read through this problem.
**Prompt for other factors:** Was there anything else that you considered when coming up with your answer?

*Figure 8. Interview prompts for question 1*

After these questions, students were given a counterargument that differed from their claims. For example, if a student identified that only fishing rods had an impact on number of fish caught, a counterargument presented to the student would be that location could also have an impact on number of fish caught given that both Tom and Jerry fished at location A, had long fishing rods, and caught the same number of fish. Students were
asked whether this changed their original claim and if so, why it changed their initial claim.

If it did not change their initial claim, students were asked what they could say to justify their original claim.

Question 2 interview questions had similar goals as question 1 interview questions, prompting for claim, evidence, warrant, process, and other factors. The questions guiding interviews on question 2 are listed in Figure 9.

- **Prompt for claim:** Which lizard did you identify as most fit?
- **Prompt for evidence:** Why do you say that?
- **Prompt for warrant:** How do you know?
- **Prompt for process:** When approaching this problem, what initially caught your attention?
- **Prompt for other factors:** Was there anything else that you considered when coming up with your answer?

*Figure 9. Interview prompts for question 2*

After these questions, students would be given a counterargument. For example, if a student responded that lizard B is the fittest because it has mated with many lizards, the implicit warrant is that fitness is defined by reproductive rate of an individual. The counterargument would target the warrant to see if students would keep the same reasoning pattern and justify/defend their warrant or follow alternate reasoning paths and either change their warrant or change their claims. Appendix D contains the interview protocol.

### 3.5 Data Analysis

Statistical analyses conducted on the LCTSR and TSA look for 1) student strengths and weaknesses, 2) correlations within sections on each assessment, and 3) correlations across sections of each assessment. Coding schemes were developed for the open response items and interviews based on existing frameworks for reasoning and arguments. Each item
on the open response questionnaire was coded by three researchers in iterative cycles that included rubric revision until Cohen’s kappa (κ) was greater than .70. Analysis processes are presented in the following sections.

### 3.5.1 Analysis of the LCTSR and TSA.

Descriptive statistics were calculated for each assessment to determine where students performed well and where students struggled. After that, correlations between each section of the LCTSR (proportional thinking, identification and control of variables, probabilistic thinking, correlational thinking, and hypothetical-deductive thinking) and each section of the TSA (claim, fact, data, opinion; qualifiers; claim or not a claim; authority, logic, theory; rebuttal, counterargument; quality of reasoning) were calculated. The parametric statistical analysis for the correlations was Pearson’s r. The results section only reports on analyses with p<.05. These analyses were conducted to explore research question 1 with the goal of understanding student skills in reasoning, their abilities to identify and understand various parts of an argument, and whether any of these skills are correlated with one another.

The results from the correlation analyses will be used to understand student responses in the open response questionnaire and interviews. If the students perform well on a particular category in reasoning or argumentation, it could provide insight into why student responses in the questionnaire and interviews exhibit certain patterns. For example, if the students score high on the claim, fact, data, or opinion section of the TSA but low on quality of reasoning, that could explain why, when asked to defend arguments, students have trouble targeting warrants. Quantitative analyses supported by qualitative data
provides a general overview of students’ abilities and then probes for underlying processes that could either support or explain findings.

### 3.5.2 Analysis of the ORQ.

A coding scheme for each question in the open response questionnaire was constructed, guided by frameworks for informal reasoning by Wu and Tsai (2007) and Sadler and Zeidler (2005). Initially, it was anticipated that students’ reasoning would fit into three broad categories: intuitive, emotive, and rationalistic reasoning. In the first round of coding, these three categories were applied to every response and item in the questionnaire. However, these categories were not specific enough to capture the small, nuanced differences in students’ responses. Additionally, some emergent reasoning patterns were identified in the questions framed in natural selection that did not fit in any of those three categories. Thus, in order to capture the full spectrum of student responses, coding schemes for each question were developed using emotive, rationalistic, and intuitive reasoning as a foundation, but also including these emergent codes.

Each item on the questionnaire contained two parts. The first part was designed to elicit a claim from students. The second part probed for the reasoning and evidence for their claim. The codes for the first part of every question takes into consideration some aspect of content knowledge, including misconceptions (Anderson, Fisher, & Norman, 2002). The codes are described in the following section. For every section, responses coded 0 correspond to unrelated or blank responses.
3.5.2.1 Coding for question 1.

The first part of question 1 prompted students to identify connections between thickness of fish hooks, length of fishing rods, location, and number of fish caught. It included data from 3 individuals all with different combinations of the possible variables that could impact the number of fish caught. A summary of the codes associated with question 1, part a is below (see Table 5).

Table 5. Codes for question 1, part a

<table>
<thead>
<tr>
<th>Code</th>
<th>Code Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Basic Claims</td>
<td>• Unrelated responses/blank responses</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• States who had the most fish without elaboration on variables (rods, location, hook)</td>
</tr>
<tr>
<td>1</td>
<td>One variable consideration</td>
<td>• Identifies rod or location (but not both) as having the biggest impact on number of fish caught OR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Only identifies that fish hooks have no relevance and does not mention either rod or location</td>
</tr>
</tbody>
</table>

Continued
The second part of question 1 asked students why they put forth that claim. The purpose of this question was to identify patterns in how students justify their answers. Codes 1 and 2 are methods of justification drawn from literature reviews (e.g., Cavagnetto, 2010; Evagorou & Osborne, 2013; Manz, 2014; Wolfe, 2011). Responses coded 1 use alternate explanations, responses coded 2 state evidence, and responses coded 3 criticize the experimental design. The criticism of experimental design was an unanticipated mode of justification, emergent from responses. A summary of the codes associated with question 1, part b is below (see Table 6).

Table 6. Codes for question 1, part b

<table>
<thead>
<tr>
<th>Code Number</th>
<th>Code Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Unrelated</td>
<td>• Unrelated or blank responses</td>
</tr>
<tr>
<td>1</td>
<td>Alternate explanations</td>
<td>• Puts forth alternate explanations to explain data OR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Puts forth different data than what is provided</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Continued</td>
</tr>
</tbody>
</table>

Table 5 continued

<table>
<thead>
<tr>
<th>Code</th>
<th>Code Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Multiple variable consideration</td>
<td>• Identifies both rod and location as having the biggest impact on number of fish caught</td>
</tr>
<tr>
<td>3</td>
<td>Experimental design</td>
<td>• Identifies that no conclusion can be drawn due to changes in variables, flawed experimental design, and/or not enough data</td>
</tr>
</tbody>
</table>
### 3.5.2.2 Coding for question 2.

In part a of question 2, students were asked which lizard was the most fit. Responses coded 0 meant the student responded that lizard a, c, or d was the most fit and responses coded 1 indicate the student responded that lizard b was the most fit. The purpose of these codes was to separate students that correctly identified the correct lizard from students that did not. For part b, students were asked how and why they came to the conclusion. Table 7 summarizes the coding scheme.

#### Table 7. Codes for question 2, part b

<table>
<thead>
<tr>
<th>Code Number</th>
<th>Code Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Unrelated</td>
<td>• Unrelated or blank response</td>
</tr>
<tr>
<td>1</td>
<td>Authority</td>
<td>• Uses support from authority (e.g., based on biologists’ definition of fitness)</td>
</tr>
</tbody>
</table>
### Table 8. Codes for question 2, part a

<table>
<thead>
<tr>
<th>Code Number</th>
<th>Code Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Unrelated</td>
<td>• Unrelated or blank responses</td>
</tr>
<tr>
<td>1</td>
<td>Naïve Mechanism</td>
<td>• Environmental pressures cause mutations and variations to arise</td>
</tr>
<tr>
<td>2</td>
<td>Teleological</td>
<td>• Mutations and/or variations can arise due to an organisms’ want or need</td>
</tr>
<tr>
<td>3</td>
<td>Mutations/sexual recombination</td>
<td>• Random mutations and/or sexual recombination is a cause for variation</td>
</tr>
</tbody>
</table>

Part b asked students how and why they came up with their responses from part a. Codes were designed based on the structures of their response, looking at whether it was...
data driven or included reasoning that explains how the claim can be drawn from the evidence. Table 9 summarizes the codes for question 3, part b.

Table 9. Codes for question 3, part b

<table>
<thead>
<tr>
<th>Code Number</th>
<th>Code Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Unrelated</td>
<td>• Unrelated or blank responses</td>
</tr>
<tr>
<td>1</td>
<td>Authority</td>
<td>• Uses generalized evidence without specificity OR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Uses support from authority</td>
</tr>
<tr>
<td>2</td>
<td>Mechanism</td>
<td>• Restates data and comments from question 2 to support what they think causes variation OR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Uses statements such as “environmental pressures” or “competition” without linking how environmental pressures or competition causes variation</td>
</tr>
<tr>
<td>3</td>
<td>Example Process</td>
<td>• Statement uses reasoning through specific examples OR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Statement describes a process for how variations arise</td>
</tr>
</tbody>
</table>

3.5.2.4 Coding for question 4.

Question 4 was similar in structure to question 3, having two parts. Part a asks students to name possible causes for one species to change into three over time and then part b asks how and why they came to that conclusion. This question was asked with the intent of seeing if there would be similarities in justification patterns when the format of the question was held to be the same, but focusing on a different concept in natural selection (origin of speciation rather than origin of variation). The emergent codes from question 4,
part a were based on common student misconceptions in speciation (Anderson et al., 2002). These codes are summarized below (see Table 10).

Table 10. Codes for question 4, part a

<table>
<thead>
<tr>
<th>Code Number</th>
<th>Code Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Unrelated</td>
<td>• Unrelated or blank responses</td>
</tr>
<tr>
<td>1</td>
<td>Naïve Mechanism</td>
<td>• Naïve theories of natural selection and genetic flow, natural disasters, cross breeding, or separation without including that speciation occurs over time</td>
</tr>
<tr>
<td>2</td>
<td>Teleological</td>
<td>• Claim reflects essentialism</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Claim is teleological</td>
</tr>
<tr>
<td>3</td>
<td>Speciation</td>
<td>• Claim is that species can change over time if a population is isolated</td>
</tr>
</tbody>
</table>

Similar to the codes for question 3, part b, the codes for question 4, part b show whether students used authority to support their claims, discuss mechanisms for speciation, or describes a process for speciation by giving an example. The codes for question 4, part b are described below (see Table 11).
Table 11. Codes for question 4, part b

<table>
<thead>
<tr>
<th>Code Number</th>
<th>Code Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Unrelated</td>
<td>• Unrelated or blank response</td>
</tr>
<tr>
<td>1</td>
<td>Authority</td>
<td>• Uses generalized evidence without specificity or support from authority</td>
</tr>
<tr>
<td>2</td>
<td>Mechanism</td>
<td>• Uses one or more of the following: environmental pressures, cross breeding, mutations, human pressures, or processes involving changes in one individual without relating it to how that causes speciation</td>
</tr>
<tr>
<td>3</td>
<td>Example Process</td>
<td>• Uses justifications contextualized through specific examples</td>
</tr>
</tbody>
</table>

3.5.3 Interview analysis.

Determining how to code the interview took multiple iterations. Initially, the analysis was done turn by turn, where each turn signifies a change in speaker. In doing so, some of the smaller nuances within statements were lost. However, looking closely at the data with a small grain size would make it difficult to see the big picture created by grouping multiple turns and coding them together. In order to find as many patterns across responses as possible, responses were coded on both levels: turn by turn for general patterns. Then, any turn that needed to be further analyzed was noted, broken down, and examined in further depth.

Interviews gave an additional perspective that informs on the results from the open responses. Students could reference their completed questionnaires during the interview and, because the interview drilled down on their responses to the questionnaire, it was possible to probe students further and parse subtle differences in responses.
Interviews based on responses to question 1 on the open response questionnaire were analyzed in two parts: claim and justification. In the claim portion, the nature of the claim was examined based on how it was presented by students, including the context in which it was put forth and the data accompanying the claim. Codes included:

- Pragmatic claims: ones that consider realistic restraints and concrete thinking, without hypothetical or theoretical considerations
- Abstract claims: ones where students reason through the data that is provided and draws a conclusion
- Coordinated claims: ones that include personal experiences when analyzing the data in the table

For the justification portion, students were given a counterargument (posed by the interviewer) to explore whether they would be persuaded towards the counterargument or defend their claims. If students defended their claims, that process could also be analyzed. These codes included:

- Experience based informal: using personal experiences to justify claims
- Evidence based formal: using classroom knowledge to justify claims
- Targeted: target weaknesses in alternate claims to justify their own claims

Interviews based on responses to question 2 were analyzed turn by turn with a code for every prompt made by the interviewer and every response made by the interviewee. These codes were used to see how students put forward claims and justified their claims when given a counterargument. When putting forward claims and justifying them, students would explicitly set grounds or a context for their claim. Students did not provide evidence or warrants for their claims unless they are prompted to do so.
Using a combination of quantitative and qualitative data helps to paint a more robust understanding of student thinking. Descriptive statistics and correlations using Pearson’s r on sub-scales of the LCTSR and TSA provided results on which skills students struggled most with and connections between skills in reasoning and skills in argumentation. Coding of the open response questionnaires revealed a pattern that was unanticipated based on literature review, indicating that there is a unique aspect to students’ claim development. Students’ responses to counterclaims in interviews show that their methods for justifying arguments are impacted by their personal experiences. These results will be discussed in further length in the following chapter.
Chapter 4: Results

In this chapter, the results of statistical analysis of responses on the Lawson Classroom Test of Scientific Reasoning (LCTSR) and the Test of Scientific Argumentation (TSA) are presented. These results include the two sub-scales on each assessment that had the highest and lowest scores. Additionally, significant correlations between skills in reasoning and argumentation are presented and discussed. Coding descriptions, counts, exemplars, and interview transcripts are examined in the subsequent sections.

4.1 LCTSR and TSA

Descriptive statistical analysis of the LCTSR and TSA showed that students scored poorly on both assessments. The mean score on the LCTSR was 10.62 out of 20 possible points (n=21) and for the TSA, the mean score was 25.43 out of 36 possible points (n=21). Table 12 summarizes the total scores on the two assessments.

\begin{table}[h]
\centering
\begin{tabular}{lcccc}
\hline
 & Minimum & Maximum & Mean & Std. Deviation \\
\hline
LCTSR Total Score & 2 & 19 & 10.62 & 4.924 \\
TSA Total Score & 15 & 35 & 25.43 & 5.537 \\
\hline
\end{tabular}
\caption{Descriptive statistics: Student scores}
\end{table}

Students may have scored low on the LCTSR because questions on this assessment are framed within science or math contexts. Because this small sample of students included many non-math and non-science majors, it is expected that students would struggle on a
scientific reasoning assessment. The TSA used general statements that were not framed within a specific science context, which could account for the higher scores on the TSA. On this assessment, students were asked to evaluate statements in day-to-day contexts rather than in science contexts. This could have been a contributing factor to the higher scores. In addition, some of the terms on the TSA (e.g., claims, authority, opinions) are terms students may encounter more frequently in everyday life, making them more familiar.

4.1.1 LCTSR Results.

Closer analysis of the components of each assessment revealed that students had the most difficulty with probabilistic thinking, with a mean score of 3 out of 10 total possible points (n=21). Students’ trouble with probabilistic thinking indicates that a weakness in their reasoning process lies in assessing the likelihood of occurrence of one claim among multiple claims based on data. This may affect their argumentation, particularly in addressing counterarguments. If students are unable to assess the viability of claims, they are more likely to be swayed by counterarguments rather than critically assessing the validity of the counterargument. Table 13 shows scores on each section of the LCTSR.
**Table 13. Sub-scales of the LCTSR**

<table>
<thead>
<tr>
<th></th>
<th>Number of Items</th>
<th>Mean Score</th>
<th>Average Percent Correct</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportional thinking</td>
<td>4</td>
<td>1.81</td>
<td>45%</td>
<td>1.692</td>
</tr>
<tr>
<td>Identification and control of variables</td>
<td>6</td>
<td>2.67</td>
<td>44%</td>
<td>2.008</td>
</tr>
<tr>
<td>Probabilistic thinking</td>
<td>10</td>
<td>3.00</td>
<td>30%</td>
<td>1.483</td>
</tr>
<tr>
<td>Correlational thinking</td>
<td>2</td>
<td>1.24</td>
<td>62%</td>
<td>.944</td>
</tr>
<tr>
<td>Hypothetical-deductive thinking</td>
<td>4</td>
<td>2.05</td>
<td>51%</td>
<td>1.024</td>
</tr>
</tbody>
</table>

Students scored highest in correlational and hypothetical-deductive thinking. These scores show that, in identifying patterns in data and assessing the likelihood of claims, students performed well. In argumentation, this indicates that students are able to sift through data to a conclusion. However, coupled with their weakness in probabilistic thinking, they may be challenged when it comes to looking through multiple explanations and weighing the merits of one against another.

**4.1.2 TSA Results.**

In sub-scales of the TSA, students scored highest on two subscales: 1) being able to identify claims, facts, data, and opinions and 2) identifying claims from non-claims with mean scores of 5.19 out of 6 total possible points on both categories (n=21). Students struggled most on distinguishing between authority, logic, and theory and identifying whether statements were rebuttals or counterarguments. The two categories with the lowest scores both had mean scores of 3.71 out of 6 total possible points. This could be due to students’ familiarity with putting forward claims and recognizing the difference between claims and evidence. However, they lacked an understanding about other viewpoints and
validity of evidence (McNeill, 2010; Lee et al., 2014, Jimenez et al., 2000). Error!

Reference source not found. gives descriptive statistics on each sub-scale of the TSA.

Table 14. Sub-scales of the TSA

<table>
<thead>
<tr>
<th>Sub-scale</th>
<th>Number of Items</th>
<th>Mean Score</th>
<th>Average Percent Correct</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claim, fact, data, opinion</td>
<td>6</td>
<td>5.19</td>
<td>86%</td>
<td>1.123</td>
</tr>
<tr>
<td>Qualifiers</td>
<td>6</td>
<td>3.86</td>
<td>64%</td>
<td>2.744</td>
</tr>
<tr>
<td>Claim or not a claim</td>
<td>6</td>
<td>5.19</td>
<td>86%</td>
<td>1.470</td>
</tr>
<tr>
<td>Authority, logic, or theory</td>
<td>6</td>
<td>3.71</td>
<td>62%</td>
<td>1.347</td>
</tr>
<tr>
<td>Rebuttal, counterargument</td>
<td>6</td>
<td>3.71</td>
<td>62%</td>
<td>2.028</td>
</tr>
<tr>
<td>Quality of reasoning</td>
<td>6</td>
<td>4.24</td>
<td>70%</td>
<td>1.670</td>
</tr>
</tbody>
</table>

Students’ scores on subscales of the TSA are to be expected, based on their results on the LCTSR. Students who performed well in 1) identifying claims from fact, data, or opinions and 2) classifying statements as claims or non-claims also performed well in hypothetical-deductive thinking. Being able to theorize claims using hypothetical-deductive thinking requires students to identify claims and view them as completely separate from data. The trouble students had with both the authority, logic, and theory section and the counterargument and rebuttal section stand to reason when looking at students’ problems with probabilistic thinking. Probabilistic thinking includes using logic to assess the validity, likelihood, or viability of multiple claims. In addition, probabilistic thinking requires assessing the grounds (reasons) supporting the acceptability of statements while also considering counterarguments.
4.1.3 Correlations across the LCTSR and TSA.

Correlations were calculated for sub-scales both within the LCTSR and TSA and across sub-scales of each assessment.

Analysis of correlations between categories of reasoning and arguments revealed two significant correlations between reasoning sub-scales on the LCTSR and argumentation sub-scales on the TSA. A positive correlation existed between identification and control of variables and determination of whether authority, logic, or theory are the basis for believing a claim ($r=.536$, $p<.05$). This may have been attributed to students’ abilities in understanding the source of evidence and how evidence fits in experimental design. For example, students who were able to identify claims supported by authority as opposed to claims supported by logic were also able to see how parts of an argument fit together in terms of summarizing an experiment. The correlation between identification and control of variables with probabilistic thinking indicates that the strength of students’ abilities to identify and manipulate variables within an experiment was a predictor of their ability to assess the viability of claims and select one from multiple. When presented with counterarguments, these students were more resistant to change because they could identify data.

The second significant positive correlation was between a sub-scale on the LCTSR containing items testing students’ hypothetical-deductive reasoning and a sub-scale on the TSA prompting students to label statements as claims and non-claims ($r=.492$, $p<.05$). These connections were reasonable because hypothetical-deductive reasoning includes developing a hypothesis and determining whether evidence from observations support or refute the hypothesis. In the practice of authentic science, the first step in hypothetical-
deductive thinking is proposing a testable hypothesis, which also includes the ability to create claims.

4.1.4 Correlations within the LCTSR and TSA.

Some other significant correlations were found within sub-scales of the LCTSR and within sub-scales of the TSA. On the LCTSR, there were moderate positive correlations between proportional and correlational thinking (r=.531, p<.05) and identification and control of variables and probabilistic thinking (r=.436, p<.05). On the TSA, the identifying statements as claim, data, fact, or opinion and labeling statements as claims or not a claim had a moderate positive correlation (r=.461, p<.05). All correlations are summarized in Error! Reference source not found.

Table 15. LCTSR and TSA Correlations

<table>
<thead>
<tr>
<th></th>
<th>Probabilistic thinking</th>
<th>Correlational thinking</th>
<th>Claim or not a Claim</th>
<th>Authority and Logic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportional</td>
<td>Pearson Correlation</td>
<td>.359</td>
<td>.531*</td>
<td>.337</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.110</td>
<td>.013</td>
<td>.135</td>
</tr>
<tr>
<td>Identification and control of</td>
<td>Pearson Correlation</td>
<td>.436*</td>
<td>.413</td>
<td>.227</td>
</tr>
<tr>
<td>variables</td>
<td>Sig. (2-tailed)</td>
<td>.048</td>
<td>.063</td>
<td>.225</td>
</tr>
<tr>
<td>Hypothetical-deductive</td>
<td>Pearson Correlation</td>
<td>.033</td>
<td>-.012</td>
<td>.492*</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.887</td>
<td>.958</td>
<td>.023</td>
</tr>
<tr>
<td>Claim, Data, Fact</td>
<td>Pearson Correlation</td>
<td>.090</td>
<td>-.139</td>
<td>.461*</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.698</td>
<td>.547</td>
<td>.035</td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.05 level (2-tailed).
4.1 Open response questionnaire results

4.2.1 Question 1a

The prompt for Question 1a is shown below. The question in italics (part a) is the focus for this section:

1. Tom, Jerry, and Dan are good friends and go fishing together most weekends. They often use the same type of fishing tools and have similar skills in fishing (that is, they often each catch a similar number of fish every time). On their last fishing trip, they had selections of different fishing rods and fishhooks, and they each picked a different location to fish (see the conditions given in the table below). They fished for a total of two hours and the number of fish they each caught during this period shown below.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Tom</th>
<th>Jerry</th>
<th>Dan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fishing rods</td>
<td>Long</td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td>Fishhooks</td>
<td>Thick</td>
<td>Thin</td>
<td>Thin</td>
</tr>
<tr>
<td>Locations</td>
<td>Point A</td>
<td>Point A</td>
<td>Point B</td>
</tr>
</tbody>
</table>

Number of fish caught during the two-hour

|          | 15 | 15 | 8 |

a) What can you say about the link between each of the three conditions and the number of fish caught?

b) Why do you say that?

The codes for this question were based on what the students identified as the variable(s) that had the biggest impact on the number of fish caught. Responses were coded
with 0, 1, 2, or 3. Table 16 presents each code with its description, the number of responses, and an exemplar student response.

Responses coded as 0, indicated that the participant did not identify any connection between variables and number of fish caught. Rather, they re-stated the information found in the table. Responses coded 1 (one variable consideration) focused only on the long fishing rod or the location as variables that impact the number of fish caught but not both. These responses may have also identified that fish hooks do not impact the number of fish caught in lieu of considerations toward fishing rod or location. These responses focused only on one variable and its relation to number of fish caught without elaborating on any other variables or considering that multiple variables could have impacted the number of fish caught. Responses coded as a 2 (multiple variable consideration) took both fishing rods and location into consideration as variables that affected the number of fish caught. These responses provided analysis of multiple variables and their relationships. Responses coded 3 were unanticipated responses in an emergent pattern. These students’ responses analyzed the changes in variables across Tom, Jerry, and Dan, identifying that there is no true control of variables across these three data points because Dan differs in two variables compared to Tom and Jerry. Dan fishes in a different location and uses a short fishing rod so the difference in number of fish caught cannot be fully attributed to any variable. The only way to truly understand the links between the three conditions and the number of fish caught would have been to add data points such as another individual with a short rod and thick hook at Point A or another individual at Point B with a long rod and thick fish hook.
Table 16. Coding and exemplars for question 1, part a

<table>
<thead>
<tr>
<th>Code Number</th>
<th>Code Name</th>
<th>Description</th>
<th>Number of responses</th>
<th>Exemplars</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Basic Claims</td>
<td>• Unrelated responses/blank responses&lt;br&gt;OR&lt;br&gt;• Responses that state who had the most fish without elaboration on variables (rods, location, hook)</td>
<td>2</td>
<td>“Tom and Jerry caught the most fish.”</td>
</tr>
<tr>
<td>1</td>
<td>One variable consideration</td>
<td>• Identifies rod or location as having the biggest impact on number of fish caught&lt;br&gt;OR&lt;br&gt;• Only identifies that fish hooks have no relevance</td>
<td>6</td>
<td>“The long fishing rods caused many fish which is contrary to the short ones.”</td>
</tr>
</tbody>
</table>
Table 16 continued

<table>
<thead>
<tr>
<th></th>
<th>Multiple variable consideration</th>
<th></th>
<th>Experimental design</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Identifies both rod and location as having the biggest impact on number of fish caught</td>
<td>10</td>
<td>Identifies that no conclusion can be drawn due to changes in variables, flawed experimental design, and/or not enough data</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>“Tom and Jerry caught same amount of fish due to having some type of long fishing rods and same point A location. No difference with fishhooks either thick or thin. Dan on the other hand had different short fishing rod and had different location. Ended up getting low count of fish.”</td>
<td>3</td>
<td>“Long fishing rods work better than short when catching fish. Fish hooks do not make a big difference. I would have Tom and Jerry change to a small fishing rod and retest theory.”</td>
<td></td>
</tr>
</tbody>
</table>
The data revealed different levels of ability among participants when they identify variables, showing that most students can identify patterns in data to formulate a claim. While some students were able to identify multiple variables, some students were not. Almost half of students only considered one variable (either fishing rods or location) as having an impact on number of fish caught. Nearly half of the students (10 out of 21) noticed that both long fishing rods and location were similarities between Tom and Jerry. Six respondents noticed one similarity (either fishing rods or location) and stopped. These students were unable to see beyond one variable to understand the importance of analyzing data and the controlling variables to reduce confounding factors in experimental design.

4.2.2 Question 1b.

The prompt for question 1b is below. The question in italics (part b) is the focus of this section.

1. Tom, Jerry, and Dan are good friends and go fishing together most weekends. They often use the same type of fishing tools and have similar skills in fishing (that is, they often each catch a similar number of fish every time). On their last fishing trip, they had selections of different fishing rods and fishhooks, and they each picked a different location to fish (see the conditions given in the table below). They fished for a total of two hours and the number of fish they each caught during this period shown below.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Tom</th>
<th>Jerry</th>
<th>Dan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fishing rods</td>
<td>Long</td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td>Fishhooks</td>
<td>Thick</td>
<td>Thin</td>
<td>Thin</td>
</tr>
<tr>
<td>Locations</td>
<td>Point A</td>
<td>Point A</td>
<td>Point B</td>
</tr>
<tr>
<td>Fish Caught</td>
<td>15</td>
<td>15</td>
<td>8</td>
</tr>
</tbody>
</table>
a) What can you say about the link between each of the three conditions and the number of fish caught?

b) Why do you say that?

Question 1b elicited students’ evidence and warrant(s) for their responses in 1a. If students put forward other variables that were not included in the data as possible explanations for the data, their response received a code of 1 (alternate explanations). Student responses that only restated the evidence present in the table received a code of 2 (statement of evidence). A code of 3 (criticism of design) signified that the student’s response paid particular attention to the experimental design and data. These responses indicated that the 3 data points (Tom, Jerry, and Dan) did not provide enough information for students to confidently put forth a claim about connections between the variables of length of fishing rods, thickness of fish hooks, location, and number of fish caught. These students described reasonable control of variables as a requirement for designing experiments and drawing inferences from data. Table 17 contains descriptions of the codes, numbers of responses, and exemplars for each code.

Responses coded with a 0 were blank or unrelated responses. Students that provided very little information or one word answers were also given a 0 code because inferences could not be made on student was thinking in these instances. In responses coded with a 1, students put forward other explanations for the results in the table. These usually resulted from students’ past fishing experiences. Code 2 responses re-stated the information give in the table as enough support for their responses in 1a. These responses restated what was given in the prompt and did not draw from any personal experiences. Code 3 responses explained why no conclusion could be drawn with any reasonable degree of confidence.
Students with these responses used their knowledge of designing experiments and controlling variables to limit confounding variables. Students identify that there were too many changes in variables across Tom, Jerry, and Dan. These students described a need for additional data points to thoroughly test the relationship between variables.

Table 17. Coding and exemplars for question 1, part b

<table>
<thead>
<tr>
<th>Code Number</th>
<th>Code Name</th>
<th>Description</th>
<th>Number of responses</th>
<th>Exemplars</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Unrelated</td>
<td>• Unrelated or blank responses</td>
<td>2</td>
<td>“Location.”</td>
</tr>
<tr>
<td>1</td>
<td>Alternate explanations</td>
<td>• Putting forth alternate explanations to explain data</td>
<td>1</td>
<td>“I say that because I believe that there might not be enough fishes staying at Point B and also the fishes might be staying more longer distance from where Dan was catching fish at. Also Tom and Jerry caught more fish because their fishing rod were longer so it could catch fish from deeper areas.”</td>
</tr>
</tbody>
</table>
The results from question 1b indicate that students primarily put forth given evidence as justification and commonly did not provide warrants for their statements. Given the results on the TSA, this finding was expected. Students who had difficulty determining whether the reasoning behind believing claims were from authority, logic, or theory also had trouble presenting warrants as reasons to connect their evidence to claims.

|   | Statement of evidence | • Present evidence (e.g., amount of fish) | “Tom and Jerry both used long fishing rods and were located in Point A. This resulted in them catching 15 fish compared to Dan’s 8.”
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Criticism of design</td>
<td>• Target experimental design (e.g., control of variables, explanation of variables, confounding factors)</td>
<td>“Tom and Jerry caught the same number of fish in the same location, using the same length fishing rods, but different fish hook thickness. Other condition variables could have affected the number of fish caught, like location or fish hook thickness. There was no scenario in which all conditions were the same except fishing location.”</td>
</tr>
</tbody>
</table>
4.2.3 Question 2a and b.

The prompt for question 2a and 2b are shown below:

“The Canary Islands are seven islands just west of the African continent. The islands gradually became colonized with life: plants, lizards, birds, etc. Three different species of lizards found on the islands are similar to one species found on the African continent (Thorpe & Brown, 1989). Because of this, scientists assume that the lizards traveled from Africa to the Canary Islands by floating on tree trunks washed out to sea” (Anderson, Fisher, & Norman, 2002, p. 974).

2. Fitness is a term often used by biologists to explain the evolutionary success of certain organisms. Below are description of four fictional female lizards.

<table>
<thead>
<tr>
<th></th>
<th>Lizard A</th>
<th>Lizard B</th>
<th>Lizard C</th>
<th>Lizard D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body length</td>
<td>20 cm</td>
<td>12 cm</td>
<td>10 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>Offspring surviving to adulthood</td>
<td>19</td>
<td>28</td>
<td>22</td>
<td>26</td>
</tr>
<tr>
<td>Age at death</td>
<td>4 years</td>
<td>5 years</td>
<td>4 years</td>
<td>6 years</td>
</tr>
<tr>
<td>Comments</td>
<td>Lizard A is very healthy, strong, and clever</td>
<td>Lizard B has mated with many lizards</td>
<td>Lizard C is dark colored and very quick</td>
<td>Lizard D has the largest territory of all the lizards</td>
</tr>
</tbody>
</table>

a) Which lizard might a biologist consider to be the “most fit”?

b) Why do you say that?
For question 2a, 11 students identified lizard b as the most fit while 10 students identified either lizard a, c, or d as most fit. Table 18 gives a description of each code, the number of responses, and an exemplar. The codes for part 2b differed in the structure of students’ justifications. These codes described the type of evidence used and whether a warrant was included. If students only cited support from authority as justification for their responses, they received a code of 1 (authority). Students who used data from the chart as evidence to justify and support their claims received a code of 2 (evidence only) and if they used the evidence from the chart and explained their reasoning (provided a warrant) by describing how the evidence brought them to their claim, they received a code of 3 (evidence and warrant).

Table 18. Coding and exemplars for question 2, part b

<table>
<thead>
<tr>
<th>Code Number</th>
<th>Code Name</th>
<th>Description</th>
<th>Number of responses</th>
<th>Exemplars</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Unrelated</td>
<td>• Unrelated or blank response</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Authority</td>
<td>• Statement with support from authority (e.g., based on biologists’ definition of fitness)</td>
<td>2</td>
<td>“because scientists believe ‘final of the fittest’ meaning those who are the most fit, survive long and have lots of offspring, will become the dominant ones.”</td>
</tr>
</tbody>
</table>

Continued
Table 18 continued

<table>
<thead>
<tr>
<th>Code</th>
<th>Type of Response</th>
<th>Evidence and Warrant Explanation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Evidence only</td>
<td>• Statement that puts forward evidence from the chart</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>Evidence and Warrant</td>
<td>• Statement that puts forward evidence from the chart and describes how the evidence can be used to support their choice</td>
<td>5</td>
</tr>
</tbody>
</table>

Code 1 responses appealed to authority or theories in science, but did not discuss how those theories applied to this specific context, their claim, or the data provided. These codes could be attributed to students’ lack of classroom content knowledge to support their responses. Code 2 responses re-stated the evidence given in the table, but did not describe how the conclusion was drawn from the evidence (did not provide warrants). In code 3 responses, students provided justification that included evidence and clearly discussed how they were able to connect the evidence to their claim. In the exemplar below, the response included the following as evidence: “because he has mated with many lizards.” The response went on to explain how that evidence relates back to fitness by stating: “fitness is defined as an organisms’ ability to survive and reproduce.” This response showed that
fitness is defined as an organisms’ ability to survive and reproduce and the data indicated that lizard b has mated with many lizards. Thus, it can be concluded that lizard b is the most fit. The results from this question showed that, while students could analyze evidence, they might not consistently put forth the connection between evidence and claims. Students viewed evidence as sufficient information to provide for both the “how” and “why” leading up to their conclusions.

**4.2.4 Questions 3 and 4, part a.**

Prompts for questions 3 and 4, part a are below:

1. a) According to the theory of natural selection, from where did the variations in body size in the three species of lizards most likely come?
2. a) What could cause one species to change into three species over time?

Questions 3 and 4 are similar because the prompt was given as a question with no data tables. The second part for both questions asked how and why students arrived at their conclusions. The intent of the second part of the question was to elicit both reasoning and justification for their students’ responses. Table 19 gives descriptions of each code, numbers of responses, and exemplars for question 3, part a. Table 20 provides information about question 4, part a.

For question 3, part a, if students identified environmental pressures as the cause of variations, they received a code of 1 (naive mechanism). These responses are a result of the alternative conception that mutations and variations are driven by environmental pressures. Responses received a code of 2 if they had the alternative conception that
mutations and variations are want- or need- driven, arising due to a specific organism’s wants or needs. These views were teleological, exemplifying essentialism. Students in this category indicated that organisms can decide to adapt based on the environment. Code 1 and 2 responses differed in that code 1 responses indicated environmental pressures as a driving force for mutations and variations while code 2 responses indicated that an individual organism’s need or want to adapt caused variations to arise. Responses coded 3 discussed random mutations and sexual recombination as causes for variations. These responses correctly identified that variations arise independently from environmental pressures or an organisms’ wants or needs to survive.

Rather, variations are due to random genetic changes or sexual recombination. While most of these mutations have no effect, some can be harmful or beneficial to the organism.

Table 19. Coding and exemplars for question 3, part a

<table>
<thead>
<tr>
<th>Code Number</th>
<th>Code Name</th>
<th>Description</th>
<th>Number of Responses</th>
<th>Exemplars</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Unrelated</td>
<td>• Unrelated or blank responses</td>
<td>5</td>
<td>“The three species of lizards most likely have come from the Canary Island on the African continent.”</td>
</tr>
</tbody>
</table>

Continued
For question 4, part a, responses coded 1 (mechanism) included naïve theories, such as Lamarckian inheritance. The alternative conception was that organisms needed to change to survive in their environment and, as a result, organisms adapted and were able to pass on these traits to their offspring. After many generations of offspring, a new species was made. A response coded with a 2 (teleological/essentialism) indicated that different species vary due to a want or a need to adapt and change. A response coded as a 3 (speciation) stated that changes from one species to a different species were a result of complete isolation of a population over a very long period of time.

<table>
<thead>
<tr>
<th></th>
<th>Naïve Mechanism</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Environmental pressures cause mutations and variations to arise</td>
<td>8</td>
<td>“the variations in body size came from each lizard’s environmental needs.”</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Mutations and/or variations can arise due to an organism’s want or need</td>
<td>3</td>
<td>“the lengths would have changed by what sizes allowed life to be easier. The lizards would have adapted to what size (length) fit them the best.”</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Random mutations and/or sexual recombination is a cause for variation</td>
<td>5</td>
<td>“the variations probably happened through the different lizards who mated and produced the new offspring.”</td>
<td></td>
</tr>
</tbody>
</table>
Of the students who fully responded to this question, all presented some form of alternative conception towards speciation. The students who provided mechanisms did not offer further descriptions of the mechanism or how they related to speciation. Others wrote that different species could arise due to one species encountering vastly different environments and, under the pressure to survive, this species changed into a different species to adapt to their surroundings.

**Table 20. Codes and exemplars for question 4, part a**

<table>
<thead>
<tr>
<th>Code Number</th>
<th>Code Name</th>
<th>Description</th>
<th>Number of responses</th>
<th>Exemplars</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Unrelated</td>
<td>• Unrelated or blank responses</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Naïve Mechanism</td>
<td>• Naïve theories of natural selection and genetic flow</td>
<td>10</td>
<td>“natural disasters or human intervention for example.”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Natural disasters</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Cross breeding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Teleological</td>
<td>• Reflect essentialism or are teleological</td>
<td>9</td>
<td>“the one species living in different places and having to adapt to their environment.”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Discuss separation, but does not include that speciation occurs over time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Speciation</td>
<td>• Claim is that species can change over time if a population is isolated</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
The high number of responses that received a code of 1 in both questions indicated that students continue to view natural selection as a driving force caused by an individual’s need to adapt to different types of environments. Students do not view evolution at the species level but rather as a controlled process that occurs at the individual level. These results are consistent with reported literature on alternative concepts in natural selection (i.e. Anderson, 2007) This weakness in content knowledge may be a cause for their responses to part b of this question.

4.2.5 Questions 3 and 4, part b.

Questions 3 and 4b asked students: “Please explain how and why you came to that conclusion.” Table 21 gives the description of each code, number of responses, and exemplars. All responses coded 0 were blank responses. Responses coded 1 (authority) used support from authority or general evidence (for example, theory of evolution or learned from class) without any additional information. These responses used appeals from outside sources. Students viewed “how” and “why” as asking the same question. Their responses drew from processes or words that were memorized by rote from classes. These students may not fully understand the content but have learned to associate words with concepts (e.g., natural selection, survival of the fittest, and Charles Darwin are all associated). Conclusions could not be made about whether these students’ responses were due to a lack of content knowledge or a lack of understanding about arguments because participants had varying levels of college level science classes. Their responses could be attributable to either a lack of content knowledge, a lack of understanding of argumentation, or both.
Student responses coded as 2 (mechanism) referred back to essentialism or teleological reasoning – that environmental pressures or competition caused individual organisms to adapt. For this code, students viewed “how” and “why” as prompts for data and restated pure data without explaining why the data was significant or how it fit within the context of the prompt. These responses described variations as an occurrence that helped organisms adapt to their environment.

Students received a code of 3 (example process) if they gave processes for how variations arise or discussed those processes by providing specific examples or experiences. An example of this code was a students’ response that discussed humans moving to Colorado and needing to adjust to the change in altitude. This response presented both a process and human example for explaining where variations arise. Students view “how” and “why” as a prompt for elaborating on their understanding. Rather than putting forward data, these students explained their thinking using examples.

These results show that students preferred to provide descriptions or examples to illustrate their points, especially in open-ended questions that did not include data used to support the reasoning process. In question 2, students were given data in the form of the table with information about 4 lizards. The results for that question showed the students’ inability to connect data and claims. Instead, nearly half of the students gave a claim and restated data given in the table. For question 3, students were not given any data. As a result, their responses included claims and examples used to explain their reasoning. The exemplar in Table 21 for code 3 (example process) shows this point.
### Table 21. Coding and exemplars for questions 3 and 4, part b

<table>
<thead>
<tr>
<th>Code Number</th>
<th>Code Name</th>
<th>Description</th>
<th>Number of responses (Q3/Q4)</th>
<th>Question 3b Exemplars</th>
<th>Question 4b Exemplars</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Unrelated</td>
<td>• Unrelated or blank responses</td>
<td>3/2</td>
<td>“I came to this conclusion because I’ve been through classes that have taught me about natural selection and natural selection choose the best features for species to live in.”</td>
<td>“once again, Charles Darwin’s theory. As they evolve they adapt to help them survive”</td>
</tr>
<tr>
<td>1</td>
<td>Authority</td>
<td>• Uses generalized evidence without specificity OR • Support from authority</td>
<td>4/2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 21 continued

<table>
<thead>
<tr>
<th>2</th>
<th>Mechanism</th>
<th>6/11</th>
<th>“In natural selection you adapt to be more fit where you are because you’ll have a better chance of survival.”</th>
<th>“because a change in food chain or a natural disaster can cause this to happen”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Restates data and comments from question 2 to support what they think causes variation OR • Uses statements such as “environmental pressures” or “competition” without linking how environmental pressures or competition causes variation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>3</th>
<th>Example Process</th>
<th>8/6</th>
<th>“The larger lizard was also the strongest one, the habitat in which it lives may require it to be stronger while the smaller ones may not need strength as much as speed or endurance. A human runner will be faster with more endurance than a body builder, yet be smaller in stature. I believe it is also a widely accepted idea that animals evolve based on their needs in their environment. Deep water fish have no need for sight, hence they do not have functional eyes.”</th>
<th>“Birds that are traveling due to migration may have to stay at an island that is unfamiliar to them, however they stay here and begin to adapt and live on that island. This is what happened to Darwin’s finches and it is the best example of how species differ based on their environment and how evolution causes the species to adapt.”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 21 continued
To explain the origin of variation, the student stated that “animals evolve based on their needs in their environment.” The evidence that this student provided was through examples, such as human runners that need endurance rather than strength or deep-water fish that do not have functional eyes. The warrant being that, because we observe this variation in species of different habitats, it must be that evolution was based on environmental needs.

4.3 Question 1 Interview Results

Patterns accompanying claim development were pragmatic claims, data-based claims, and coordinated claims. Pragmatic (rationalistic) claims were ones put forward framed within a specific context. These took into consideration realistic restraints without thinking about the theoretical or broad picture. These responses tended to weigh more heavily on personal experiences than classroom or science knowledge. In data-based claims, students assessed the data presented and summarized the data in a way that supported their claims. These were inductive or deductive. Coordinated claims were ones that considered both personal experiences and classroom knowledge. These responses were framed in terms of personal experiences. Students that fell into this category put forth a combination of classroom knowledge and personal experiences as the claim.

In looking at how students justify their claims in this decontextualized scenario, patterns that emerged were experience-based informal, which only used personal experience as evidence. Justifications that used what is learned in class as evidence were classified as evidence based formal.
For question 1, there were 5 students who put forward data-based claims, 4 students with pragmatic claims, and 1 student who had coordinated claims. The pragmatic claims used evidence based informal justifications and abstract claims used evidence based formal justifications. Due to time constraints, the one student who had a coordinated claim was not able to complete the interview. Thus, that student’s justification will not be presented.

4.3.1 Interview 1: Pragmatic claim with experience based informal justification.

Interview Transcript Exemplar 1: Pragmatic, experienced based

Interviewer-So let’s go ahead and move on to question 1. What did you think had impact on the number of fish caught? (claim prompt)

Student- The most impact? Or just—I’m thinking everything could have. But being a fisherman and fishing a ton, I would say it’s the location. (claim)

I-Why do you say that? (evidence prompt)

S-because my experience in this is that location will have an impact. It’s not so much luck when you’re fishing, the fishing rod could make a difference of how it affects the casting and reeling back, but the major thing is the location. That’s why when you’re fishing, you change to different places. (evidence – pragmatic reasoning)

I-So when you were coming up with your decision about it, did you look at the information in the table at all? (process prompt)

S-I made my assumptions based on my experience and then looking at the table, I knew. But I think it had to do with my experience more. You can see in point a and point a, they got 15 but Dan was fishing in point b. maybe he should have been fishing in point a (evidence-experience; use evidence in the table to support experience)

I-you’ll see tom and jerry both fished in point a and they caught 15 fish and they both used long fishing rods, but you still felt location had more to do with fish caught? (counterargument/probe for justification)
S—yeah, in my experience, with the shorter rod, it’s not the fact that you can’t really—it may be somewhat you can’t cast as far but it doesn’t say where they were catching the fish. Out far or by the boat or out at the shore? So the short rod gives you action in reeling the fish better but that’s what makes me still say it’s location. (experience based)

In this interview exchange, the student indicated that the location was the variable with the highest impact due to his or her experience in fishing. Even after being prompted about the process, the student indicated that experiences had the biggest impact on how the claim was developed When faced with a counter argument—that Tom and Jerry both also used long fishing rods, the student still felt that fishing rods would not make a difference because fishing rods would be selected based on location.

4.3.2 Interview 2: Data-based claims with evidence based informal justification.

Interview Transcript Exemplar 2: Data based, formal

Interviewer-Okay, so what connections did you say there were between the fishing rods, fish hooks, and location? (claim prompt)

Student-So there’s two different locations and two men at the location and one at the second. But--- and the same amount of fish caught at point a. My thought process was that you had both long rods but you had two different types of hooks so in my mind, that eliminated, um, the hook as a variable that separated the two and what I wrote down was just basically that point a was a “better” spot but the only variable that remains would be the length of the rod. (claim + evidence)

I-Okay. So, which one was the most influential factor—fishing rods or location or both? (clarify)

S-I was going to say location.
I-Okay. So, if I were to say… I think that thin fish hooks mean you don’t catch as much fish because Dan had a thin fish hook and he only caught eight fish and it could just be that point a has more fish which is why Jerry happened to have a thin fish hook and more so I would say that fish hooks do matter. Could you persuade me that it doesn’t? (counterargument)

S-Yeah. I would say just because Tom also had a—he had a thick hook at the same location and Jerry and Tom caught the exact same number of fish at the same spot. (evidence formal)

In this abstract reasoning example, the student put forth a claim and justification that were based only on the data in the table. This student did not take into consideration any background or outside knowledge. Rather, the student viewed the data as sufficient to convey a claim. This differs from the pragmatic claim because the pragmatic claim places the data from the table into personal experiences. The personal experiences influence how the data is interpreted.

4.3.3 Interview 3: Coordinated claims.

Interview Transcript Exemplar 3: Coordinated

Interviewer- Great. So, go ahead and read question one.

Student- Maybe location A might have had more fish and it might have been bigger and deeper fish. ‘cause the long fishing rods seem to yield more fish. Maybe they were deeper or location A had more fish. (claim)

I-Okay, so to you, fishing rods and locations were the two factors that impacted how many fish were caught? (clarify)

S-right. Because Tom and Jerry were in the same location and had the same length but they had different thickness of fishhooks. (evidence)

I-Okay.

S-And still had the same number of fish. (evidence)
I-Okay. So can you talk me through a little bit about what you were thinking before you started talking? (process prompt)

S-Um, well, the first thing that seemed relevant was the place they were fishing. Since Tom and Jerry were both at fishing point A and they both caught 15 fish, maybe there were more fish in that pond. I also noticed that both Tom and Jerry were fishing with long fishing rods so maybe there were bigger or deeper fish. Uh, but, the—Tom and Jerry had different….Tom had a thick fishing hook and Jerry had a thin one, but they caught the same number of fish even though they were….so they’re fishing in the same spot. They have the same kind of rods but they have different fish hooks so I’m thinking the kind of fish hook is irrelevant. (evidence + experience)

In this interview excerpt, the student’s claim was considered coordinated because, after the process prompt, this student took into consideration the data that was in the table, stating the location where Tom and Jerry fished and the number of fish they caught. However, this student also contextualized the data of the fishing rods based on their personal knowledge about fishing. Particularly, that the fishing rods and fish hook thickness are related to the size of the fish Tom and Jerry were trying to catch and the depth of the water at their location. This level of reasoning differs from either pragmatic or data-based because it carries aspects of both types.

4.4 Question 2 Interview Results

In interview question 2, students were asked to select the lizard that was most fit out of 4 lizards from the Canary Islands. The most frequently observed pattern occurred before students gave a claim or provided justification. Students would first try to set grounds or a context. For this interview question, students provided a definition of fitness first and did not give any evidence for their claims until prompted. Warrants also had to be explicitly probed and questioned before students provided “how” or “why” they interpreted
the evidence that led to their claim. The following interview excerpt is an example of students’ setting grounds before putting forth claims.

Question 2 Interview Transcript Exemplar: Setting Grounds

I: So what lizard would you deem to be the most fit? (claim prompt)

S: That's another tough one. I guess it depends on what my definition of fitness is. (setting grounds)

I: So what's your definition of fitness? (prompt for grounds)

S: I would think, from a personal standpoint, fitness, you live longer, and most people think about it, you wanna live longer. So I guess that would be Lizard D, living six years. (grounds)

I: Why are you shrugging?

S: I don't know. There's a lot of things to do with the term "fitness". So, I don't know, maybe I'm thinking too much into it again, but...

I: Also talk to me what's happening in your thinking process. (prompt for process)

S: Well, for example, if you mate with more lizards, the population of lizards is more likely to expand, you got more offspring surviving to adulthood with Lizard B (evidence), that's clearly mating with more lizards. So as far as species health, that seems to me like that would be the species that would propagate the most, and overall, for the species, be more healthy, as opposed to just an average lizard living longer than anything else. (change in grounds)

In this excerpt, the student wanted to define fitness before deciding which lizard is the most fit. After defining fitness as the length of life (the longer an organism lives, the more fit it is), the student put forward the claim that Lizard D is the most fit given that it lived the longest in comparison to the other lizards. However, through the process prompt, the student changed the grounds for being healthy. After talking through the evidence, the student explained that species health is the ability to propagate, rather than just the lifespan.
of a single lizard. When students had trouble setting grounds, they relied on intuitive reasoning. These students were unable to provide support for their responses even when probed.

Question 2 Interview Transcript Exemplar: Intuitive Reasoning

I-Great. Let’s move on to question number 2. This is the one with the four lizards and information about the lizards. Which one did you think would be the most fit? (prompt for claim)

S-I went back and forth on this one. I ultimately decided on lizard A (claim)

I-Why’s that? (prompt for evidence)

S-Because it has the greatest overall body length and it was also commented as being “healthy, strong and clever” whereas the other ones were more just kind of physical description versus kind of, like—it wasn’t getting characteristics of the other one in, like, healthy, or strong or smart. The other ones just kind of describe their appearance. (evidence)

I-Gotcha. Okay. So, when you decided on lizard a, you said that you were kind of torn.

S-Yeah.

I-Which other lizard were you considering? (process prompt)

S-Originally, I said b which I wouldn’t even say I settled on. B was the one I initially picked and that was because it had the greatest number of offspring to survive out of all of them and its age was median of the others (evidence). So it’s kind of, like, a…you know…I don’t know. That’s just kind of what seemed to make sense to me. (intuitive reasoning) As I thought about it, the comment said lizard b had mated with many lizards. That doesn’t seem to do a lot.

I-So you placed more emphasis on the comments. Does that sound right?

S-Yeah, I guess so. Ultimately, I did.

I-If I were to say—in my opinion, lizard d is the most fit because I think that it has lived for 6 years so that means it has to be strong or at least have high fitness level. The fact that it has the largest territory means that it could probably assert
dominance and fight off other lizards to be able to not only live that long but also have the largest territory. Could I convince you that lizard d is more fit? (counterargument)

S-Potentially, I guess it just kind of more so was based on what I was picturing should be fit. I guess I kind of imagined more of a, more of its physical characteristics. So healthy, strong, and clever came out as being what I would describe as fit versus being successful or whatever other term you’d use for that. I think I could definitely be convinced that lizard d could be it. (setting grounds)

This interview excerpt shows how this student justified his or her response in light of a counter argument. Here, the student framed their response by setting the grounds under which the claim would be correct. Since the student defined fitness as “being healthy, strong, and clever” rather than in terms of reproductive success, they did not yield to counter arguments.

Results from this chapter show that students in this sample also have the similar alternative conceptions about natural selection as high school and college-level students (Anderson, 2007). Additionally, their arguments rely heavily on personal experiences to support arguments. When creating an argument, students put forward claim and evidence as the only parts needed for a complete argument. These students’ arguments are resistant to change and, when responding to counterclaims on questions framed in natural selection, students’ responses are to target weaknesses in the argument by illustrating processes through specific examples. When that technique for justifying arguments is met with additional evidence supporting the counterclaim, students tended to shift the grounds or premise of their argument in order to account for both their claim and the counterclaim.

Based on their scores on the LCTSR and TSA, these patterns were expected. Because students struggled most with probabilistic thinking and determining whether support for an
argument is from an appeal to authority, logic, or theory, they are more likely to target evidence than negotiate through multiple explanations and analysis of warrants. These results will be further discussed in the next chapter, which will present the findings of this study as well as implications for future research.
Chapter 5: Discussion and Implications

This chapter will first discuss the findings for each research question followed by limitations of the study, contributions, and implication for future studies.

5.1 Research Question 1

Research question 1: How do community college students perform in reasoning and understanding the parts of an argument and how do these skills correlate? This research question was explored through statistical analysis of responses on the Lawson Classroom Test of Scientific Reasoning (LCTSR) and the Test of Scientific Argumentation (TSA).

5.1.1 Finding 1: Students’ strengths and weaknesses in reasoning and argumentation.

The first finding associated with this research question was that students struggled on both assessments. The mean score for the LCTSR was 10 out of 20 points (n=21) while the mean score for the TSA was 25 out of 36 points (n=21). These scores indicate that students have difficulty with in-depth reasoning and complex argumentation.

Reasoning skills vary in level of difficulty (Piekny & Maehler 2013). Students’ scientific reasoning when investigating data containing two variables is a basic competency that is easily undertaken by students, but when patterns in data are less obvious, students
have difficulty drawing conclusions (Kanari & Millar, 2004). Fundamental abilities in scientific reasoning involves hypothetical-deductive thinking and correlational thinking (Lawson, 2000). At a young age, children become adept at empirical-inductive thinking which is usage of language and experiences in a feedback loop to drawing generalized conclusions in the natural world (Lawson, Lawson, & Lawson, 1984). After mastery of empirical-inductive thinking, the next level of reasoning is hypothetical-deductive thinking, which involves an internal testing of explanations to develop a reasoned decision or method for problem solving (Lawson et al., 1984). Correlational thinking is finding patterns in data to identify relationships between variables. On the LCTSR, students performed best on hypothetical-deductive thinking and correlational thinking. This indicates that students in this sample did not have trouble developing methods for problem solving or finding general patterns in uncomplicated data. Participating in fundamental abilities of scientific reasoning are not as challenging to these students as higher levels of reasoning. Students’ scores on probabilistic thinking and identification and control of variables signify their weakness in higher levels of scientific reasoning.

On the LCTSR, students struggled most with probabilistic thinking and identification and control of variables. Probabilistic thinking is a practice used in daily encounters that includes relying on heuristics to make judgements about likelihoods of scenarios (Johnson-Laird, 1994). However, when applied to science contexts, this process can lead to logical fallacies (Johnson-Laird, 1994). In science, probabilistic thinking includes analyzing explanations for phenomena, weighing the variables associated, examining evidence and observations, and determining which explanation is most feasible. In large part, probabilistic thinking in science contexts comprises of inductive/deductive
thinking and a high level of comfort with uncertainty (Amir & Williams, 1999; Wright, Phillips, Whalley, Choo, Ng, Tan, & Wisudha, 1978). Identification and control of variables is a reasoning skill that relates to scientific ways of thinking because it requires an understanding of experimental design. The ability to identify and control variables includes knowledge about experiments such as how and why they are conducted. Students’ low scores on these sub-scales show that, while they are able to use scientific reasoning at a basic level, they are unable to apply scientific reasoning to more complex, higher level problems.

On the TSA, students scored well in identifying claims and differentiating between claim, fact, data, and opinion. Students’ scores on the TSA were expected based on findings on the LCTSR. Results on the LCTSR point to students’ strength in hypothetical-deductive thinking. The ability to develop a testable hypothesis is similar to the ability to create claims. This sample of students scored high in both categories which shows that their weakness was not in stating or identifying claims. However, students struggled most on differentiating between rebuttals and counterarguments and determining whether support for a statement relies on logic, authority, or theory. This aligns with the results on the LCTSR. Students who had trouble engaging in probabilistic thinking had difficulty countering arguments and assessing the viability of statements when multiple statements were presented. Struggling with this way of thinking led to problems with assessing arguments. Results from analysis of the TSA indicated that students also struggled with higher levels of argumentation. Together, the findings based on scores on the LCTSR and the TSA are 1) students in this sample had difficulty with complex reasoning processes and 2) students struggled with complex components of argumentation. This finding was also
supported by the interview data. All interviewed students had no trouble putting forward claims or claims with evidence without additional prompting. However, they struggled with identifying and providing warrants.

5.1.2 Finding 2: Correlations across sub-scales with the lowest scores.

The moderate positive correlation between identification and control of variables and probabilistic thinking ($r=.436, p<.05$) indicates that students who were able to understand relationships between variables in an experiment also performed well in assessing likelihoods of particular outcomes in problems. Overall, on the LCTSR, students struggled most on these two sub-scales. Their correlation shows that curricular support on one of these skills could prompt growth in the other. Stronger abilities in controlling variables and analyzing connections across variables (including their weight on the final claim) could increase students’ probabilistic thinking abilities. This finding affects current knowledge about this particular population’s scientific ways of thinking. While these students lacked content knowledge, the underlying problem was their lack of complex reasoning skills.

The positive correlation between students’ analysis of reasoning for supporting an argument and identification and control of variables ($r=.536, p<.05$) indicates that students’ problems in reasoning affected their argumentation. Identification and control of variables was a sub-scale on the LCTSR in which students struggled most. Authority, logic, and opinion was also a low-scoring sub-scale on the TSA. These results indicate that, while students had trouble with both skills, students who had a better grasp on variables in experiments were better at determining the type of warrant used to support arguments.
5.1.3 Finding 3: Correlations across sub-scales with the highest scores.

The ability to label statements as claims and non-claims was positively correlated with hypothetical-deductive thinking \((r=0.492, p<0.05)\) and identifying statements as claim, data, opinion, or fact \((r=0.461, p<0.05)\). Embedded in the skill of hypothetical-deductive thinking is devising a testable hypothesis that helps to describe or explain phenomena. Thus, the correlation between hypothetical-deductive thinking and ability to identify claims from non-claims was expected. Likewise, understanding the difference between data, fact, opinion, and claim was expected to have a positive correlation with identifying claim from non-claims. These correlations support the conclusion that this sample of students performed well in basic levels of reasoning and argumentation.

5.2 Research Question 2

Research question 2: What kinds of claims do community college students use when developing arguments? This research question was explored through the Open Response Questionnaire (ORQ) and interviews.

5.2.1 Finding 1: Written responses.

The coding results on the ORQ and interviews indicated that alternative conceptions persisted in this sample of students. Their claims contained misconceptions about natural selection, evolution, and adaptations. On questions that were not framed within a scientific setting, students claim development drew from a multitude of sources. To develop claims, they relied on personal experiences and encounters outside of the classroom to use as evidence for their claims. These responses were more robust.
scoring low on identification and control of variables on the LCTSR, some students were able to uncover flaws in experimental design when it was framed in the context of three friends going fishing. In that question, nearly half of students considered multiple variables (fishing rod and location) as having an impact on the number of fish caught. By comparison, on questions framed in science settings such as natural selection, origin of variation, and origin of species, students relied heavily on the data that was provided and less on personal experience. These findings may be attributed to a lack of content knowledge about processes in natural selection given that the majority of students in this sample have taken less than two science classes prior to this study. Their reliance on data to draw conclusions could be a result of their lack of outside knowledge to support their claims.

These claims also contained specific examples to help convey the process presented in the claim. Explicit warrants were rarely included in the written responses. The combination of reliance on personal experiences and specific examples to illustrate their claims indicate that this sample of students were concrete thinkers who struggled with abstract reasoning. Results from this section show that these students had difficulty expressing their claims with careful consideration of data in a logical argument structure that included clear reasoning to support their claims.

5.2.2 Finding 2: Interviews.

Like their written responses, students relied on personal experiences to elaborate on their claims when questions were not framed in science settings. When answering the prompt for a claim on questions framed around natural selection, students rarely gave
responses that included evidence. However, nearly all students who were interviewed set grounds before their claims. They first provided conditions under which the claims held true. Grounds included specifying their definition of science concepts before responding to questions or describing how a “scientist” might define science concepts and then indicate that their own definition was different. Results in claim development from the interviews were expected. Because students performed well in identifying claims on the TSA, they were not expected to have problems developing claims of their own. After identifying grounds for their claims, students did not have trouble developing claims. However, they did have weaknesses in Justifying their claims against counterarguments.

5.3 Research Question 3

Research question 3: How do community college students justify and/or defend their claims when presented with a counterargument? This interview portion was designed to investigate this research question.

5.3.1 Finding 1.

Interview analysis focused on one question framed in a science context and one that was not. Responses to counterarguments on the question about Tom, Dan, and Jerry fishing were more resistant to change than questions about natural selection. When presented with counterarguments about variables that had the most impact on the number of fish caught, students focused on evidence, citing their own experiences fishing or their knowledge of fishing as justification for their claims. These claims were so resistant to change that, even when presented with quantitative data that contradicts their claims, students still would not
alter their initial claims. This argumentation pattern shows the significance of connections to real world experiences in claim development and justification. For this sample, robust personal experiences were heavily relied upon when defending claims in spite of data. By comparison, the interview question that asked students to identify the most fit lizard elicited a different response and argumentation pattern.

When putting forward claims about the most fit lizard, students were less confident in their arguments. In lieu of responding directly to the prompt for claim by stating the lizard they felt was most fit, students first set grounds for their argument. They put forth the circumstances and parameters under which their claims would hold true. This response pattern may have been attributed to their lack of content knowledge. However, when given a counter argument, rather than analyzing the evidence presented or the viability and the strength of the warrant, students changed the grounds of their claim. This change is an indication of the fluidity of students’ alternative conceptions when students lack an understanding of the structure of arguments, their discourse around argumentation has an unexpected pattern. These claims were not as resistant to change as claims tied to personal experiences. If students understood more about arguments and their learned content knowledge connected more to personal experience, both their argumentation and reasoning process would grow.

5.4 Summary and Implications

The results from this study indicate that there are signification relationships across reasoning, argumentation, content knowledge, and personal experiences. Each construct has an important impact on the others. This sample of students still have problems engaging
in higher levels of reasoning and argumentation. Their difficulties in argumentation may be attributed to their lack of understanding of higher levels of argumentation, shown by their low scores in differentiating between authority, logic, and theory and rebuttals and counterarguments. However, their struggles with higher levels of reasoning may be due to a lack of content knowledge. When content knowledge is not required to respond to a question, students readily rely on their personal experiences to reason through a problem. Their claims are so deeply entrenched in personal experiences that, even despite contradictory data, students are confident enough in their reasoning to be resilient to claims. By comparison, when content knowledge is needed to respond to a question, students reason by utilizing data put forth in the question. They do not include personal experiences in these reasoning processes. The resistance to change in science content-based questions comes from students’ fluidity in setting grounds.

Past studies have emphasized the importance of distinguishing between the process of doing science and the content of science (Ryu & Sandoval, 2012). Students view the practice of science as inherently different from classroom science, which would be a contributing factor to their need to set grounds when developing claims. The students in this study not only draw distinctions between scientists and themselves, but also between scientific definitions and their own definitions of concepts in science. This aligns with current research about the need to shift science learning away from declarative knowledge and towards dialectic, social construction in order to change students’ understanding of the epistemology and nature of science (Christodolou & Osborne, 2012; Cavagnetto, 2010; Jimenez-Aleixandre et al., 2000; Evagorou & Osborne, 2013; Driver et al., 1996;). Despite
students’ lack of experience with argumentation, results from the interview indicate that students view of grounds align with Toulmin’s (1984) definition of grounds in science.

Studies have shown that the areas in argumentation that are of greatest difficulty are providing relevant evidence, using evidence to justify claims, and rebutting counterarguments (Berland & Reiser, 2009; McNeill, 2011; Osborne et al., 2004). While those studies were conducted on middle or high school students, findings from this study demonstrate that these issues persist into this population of students. Students in this sample still struggle with argumentation, despite having more credit hours of education. This aligns with suggestions from current research stating that a way to address weakness in students’ understanding of argumentation is to teach argumentation as part of the curriculum (Carles et al., 2008; Simon et al., 2004).

While there are many limitations to the study, the findings indicate that these students may be able to articulate knowledge through argumentation if given the opportunity to learn through real world, concrete experiences, connecting learned content to real world encounters. Presenting content knowledge through experiences that personally connect to students will increase their understanding of argumentation (such as putting forth data and justifying claims) and reasoning skills (such as critical analysis and selection of explanations and comfort with uncertainty).

5.5 Limitations

The limitations to this study are 1) generalizability and 2) reliability. The small sample size makes it difficult to generalize these results to the population. The coding schemes for each portion of the analysis were designed from a combination of review of
frameworks in literature and emergent patterns from the study. Three coders analyzed the open response and interview data, coding in an iterative cycle, adjusting clarity of codes until reaching agreement measured by Cohen’s kappa of 0.75. However, these codes, while applicable to this sample, may not be representative of patterns in all cases of the population. The coding scheme was guided by frameworks from literature and studies applying to other populations (e.g., undergraduates at a 4-year institution or high school students). The amount of variation in the population (from age range, experiences in science courses to reasons for taking courses at the institution) indicates that these results may not always hold true.

Because this study was a descriptive study and did not include interventions, the results provide information on the “status quo” of the sample. Further research with a larger sample will be needed to draw generalizable conclusions about the population.

5.6 Further Research

Based on the results of this study, further research needs to be conducted. A possible intervention is to build argumentation into curriculum. Helping students become familiar with identifying reasoning and understanding justification for arguments, should positively impact their ability to critically analyze arguments. By framing science content in the lens of argumentation and increasing students’ comfort with multiple explanations, students should experience an increase in content knowledge and reasoning abilities. An example intervention is to develop a unit about the origin of variation by giving primary sources such as Darwin’s notes from the Galapagos Islands. Ask students to come up with a claim based on evidence and an experiment to test that claim. Prompt students to compare
claims and justify their own. A pre-and post-assessment on content knowledge and argumentation given to this group compared to a group of students who did not participate in the lesson will explore changes in content knowledge and argumentation as a result of incorporating argumentation lessons. This intervention targets two major findings from this study. The first is that students’ reasoning skills and argumentation are closely related. Teaching argumentation should impact students’ probabilistic reasoning. Additionally, teaching content by engaging students in first person, authentic science experiences will make students more confident in their content knowledge and encourage them to participate in higher levels of argumentation.
References


Chenail, R. J. (2011). Interviewing the investigator: Strategies for addressing instrumentation and researcher bias concerns in qualitative research. The Qualitative Report, 16(1), 255.

productive thinking. *Journal of research in Science Teaching, 44*(6), 815-843.


Hill, A. G. (2016). *Predictors of Adult Community College Students' Intent to Transfer From the Community College with the Associate in Science, Associate in Applied Science, or Associate in Arts Degree to a Public University in Florida* (Doctoral dissertation, University of South Florida).


Lawson, A. E., Clark, B., Cramer-Meldrum, E., Falconer, K. A., Sequist, J. M., & Kwon,


Moore, J. C., & Rubbo, L. J. (2012). Scientific reasoning abilities of nonscience majors in


Argumentation in science. New York: Macmillan Publishing


Appendix A. Test of scientific Argumentation

These questions are all about science. They aren’t questions to find out how much people know, but they are questions about the way people talk and write when they are being scientists. For each set of questions, please define the science words first, then use these definitions and follow all the directions when answering the questions.

Definitions

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>claim</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>fact</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>opinion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In science, statements can be claims, facts, opinions or data. For each of the statements below, circle whether it is a claim, fact, opinion or data. Important: For this test, you don’t need to know whether a statement is actually true; just decide if the statement is stated as a fact, claim, opinion, or data.

<table>
<thead>
<tr>
<th></th>
<th>claim</th>
<th>fact</th>
<th>opinion</th>
<th>data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sound is a mechanical wave.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. <em>Colgate</em> toothpaste will increase enamel density.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. A diet high in whole grains will lead to a healthier heart.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4. Gravity pulls objects towards the center of mass.  
   claim  fact  opinion  data

5. A recent typhoon in the Philippines had wind speeds as high as 235 miles per hour.  
   claim  fact  opinion  data

6. I believe teenage drivers should not be allowed to use cell phones.  
   claim  fact  opinion  data

**Definition**

In each statement below, circle the qualifier. Be sure to circle only the word or short phrase that is the qualifier, not the words around it.

<table>
<thead>
<tr>
<th>Number</th>
<th>Statement</th>
<th>Qualifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>The recent changes in climate are probably due to humans’ use of carbon-based fuels.</td>
<td>probably</td>
</tr>
<tr>
<td>8</td>
<td>Almost all obese teenagers are sleep deprived.</td>
<td>all</td>
</tr>
<tr>
<td>9</td>
<td>Dumping medical waste into rivers can sometimes lead to gender imbalance in frogs.</td>
<td>sometimes</td>
</tr>
<tr>
<td>10</td>
<td>Some frogs will change their sex when they are placed in a single-sex population.</td>
<td>some</td>
</tr>
<tr>
<td>11</td>
<td>Some dogs make good hunters.</td>
<td>some</td>
</tr>
<tr>
<td>12</td>
<td>The removal of topsoil usually doesn’t allow for successful farming.</td>
<td>usually</td>
</tr>
</tbody>
</table>

For each of the statements below, circle whether it is a claim or not a claim.

<table>
<thead>
<tr>
<th>Number</th>
<th>Statement</th>
<th>Claim</th>
<th>Not a Claim</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>Mars orbits the sun in 687 days.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>The hardest element is carbon in the form of a diamond.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Students who study more tend to get higher grades.</td>
<td>Claim Not a Claim</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. At birth, the human body contains several billion cells.</td>
<td>Claim Not a Claim</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Neon atoms contain ten protons and ten electrons.</td>
<td>Claim Not a Claim</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Listening to classical music helps preschoolers learn more quickly.</td>
<td>Claim Not a Claim</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Definitions**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>authority</td>
<td></td>
</tr>
<tr>
<td>logic</td>
<td></td>
</tr>
<tr>
<td>theory</td>
<td></td>
</tr>
</tbody>
</table>

For each set of statements below, circle whether the scientist believes the claim because of **authority, logic** or **theory**. The scientist might believe something for many different reasons, but which reason does the scientist give?

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>19. Wetlands are a necessary part of many environments. I believe this because wetlands support different types of organisms, provide a place for migrating birds to rest and feed, and provide water for the animals living in the surrounding area to come and drink.</td>
<td>authority logic theory</td>
</tr>
<tr>
<td>20. Video games condition children to violence and cause them to act more violently in real life. I believe this because according to the <em>American Psychological Association</em>, children who are overexposed to violent video games are</td>
<td>authority logic theory</td>
</tr>
</tbody>
</table>
more likely to develop violent tendencies than those who are not.

21. Eating genetically modified organisms may cause diseases. I believe this because the World Health Organization cites the possibility of gene transfer from genetically modified organisms to the digestive tract or intestinal flora of humans as a possible health concern.

22. The use of a tablet computer reduces the risk of carpal tunnel syndrome. I believe this because I read it in an editorial in the *Journal of Medicine*.

23. People that eat the recommended amount of fiber are at less risk of heart disease. I believe this because cholesterol can clog arteries and lead to heart disease. It is thought that soluble fiber can soak up cholesterol, allowing the body to get rid of some of it.

24. Electric cars are more dangerous than gasoline-powered cars. I believe this because electric cars are built of lighter-weight material. Lighter-weight materials do not protect the driver well in accidents. That lack of protection will lead to more injuries.

<table>
<thead>
<tr>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rebuttal</td>
</tr>
<tr>
<td>Counter-Argument</td>
</tr>
</tbody>
</table>

| 21. Eating genetically modified organisms may cause diseases. I believe this because the World Health Organization cites the possibility of gene transfer from genetically modified organisms to the digestive tract or intestinal flora of humans as a possible health concern. | authority logic theory |
| 22. The use of a tablet computer reduces the risk of carpal tunnel syndrome. I believe this because I read it in an editorial in the *Journal of Medicine*. | authority logic theory |
| 23. People that eat the recommended amount of fiber are at less risk of heart disease. I believe this because cholesterol can clog arteries and lead to heart disease. It is thought that soluble fiber can soak up cholesterol, allowing the body to get rid of some of it. | authority logic theory |
| 24. Electric cars are more dangerous than gasoline-powered cars. I believe this because electric cars are built of lighter-weight material. Lighter-weight materials do not protect the driver well in accidents. That lack of protection will lead to more injuries. | authority logic theory |
For each pair of statements below, there is a claim and a response. Circle whether each response is a **rebuttal** or a **counter-argument**.

<table>
<thead>
<tr>
<th>Claim</th>
<th>Response</th>
<th>Rebuttal</th>
<th>Counter-Argument</th>
</tr>
</thead>
<tbody>
<tr>
<td>25. <strong>Claim</strong>: The latest flu outbreak will cause economic problems because of the cost of the vaccine.</td>
<td><em>Vaccines can be distributed very inexpensively.</em></td>
<td>rebuttal</td>
<td>counter-argument</td>
</tr>
<tr>
<td>26. <strong>Claim</strong>: Transplanting donor brain cells will repair traumatic brain injuries.</td>
<td><em>Intensive cognitive training has better results than transplanting</em></td>
<td>rebuttal</td>
<td>counter-argument</td>
</tr>
<tr>
<td>27. <strong>Claim</strong>: Objects always fall at a rate of acceleration equal to 9.8 meters/second/second.</td>
<td><em>A skydiver with a parachute falls slower than that.</em></td>
<td>rebuttal</td>
<td>counter-argument</td>
</tr>
<tr>
<td>28. <strong>Claim</strong>: Soil, light, water, and air are required for plants to grow.</td>
<td><em>Plants grow in the deep ocean where there is little light.</em></td>
<td>rebuttal</td>
<td>counter-argument</td>
</tr>
<tr>
<td>29. <strong>Claim</strong>: The four seasons are caused by the change in distance from the earth to the sun during earth’s orbit around the sun.</td>
<td><em>It is the tilt of the earth on its axis as it orbits the sun which</em></td>
<td>rebuttal</td>
<td>counter-argument</td>
</tr>
<tr>
<td>30. <strong>Claim</strong>: Heavy objects sink in water.</td>
<td><em>Ships weigh many tons and they float.</em></td>
<td>rebuttal</td>
<td>counter-argument</td>
</tr>
</tbody>
</table>

**Definition**

<table>
<thead>
<tr>
<th>Quality of Reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clai</strong>erer**</td>
</tr>
<tr>
<td><strong>Reasoning</strong></td>
</tr>
<tr>
<td><strong>Quality</strong></td>
</tr>
<tr>
<td><strong>of</strong></td>
</tr>
<tr>
<td><strong>Reasoning</strong></td>
</tr>
</tbody>
</table>
For each chain of reasoning, indicate whether the quality of reasoning is **strong** or **weak**. Use your best judgment.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studies have shown that parents and children get on each other’s nerves more as the children get older. These studies surveyed thousands of parents and their children. Therefore, short conversations, particularly between mothers and daughters, should replace longer conversations.</td>
<td>Strong</td>
</tr>
<tr>
<td>The velocity of a rolling tennis ball gradually decreases. Newton’s theory says that an object in motion stays in motion with the same speed unless acted upon by an unbalanced force. Friction is a force. So, it is probably friction that slows down the tennis ball.</td>
<td>Strong</td>
</tr>
<tr>
<td>Mrs. Washington’s class worked in three groups to test how fertilizer affects plant growth. Each group planted 10 plants in containers. One used no fertilizer, one used a small amount of fertilizer, and one used a lot of fertilizer. They found that those plants with a small amount of fertilizer grew biggest. So, the class concluded that fertilizer containing iron worked better than fertilizer containing nitrogen.</td>
<td>Strong</td>
</tr>
<tr>
<td>Kelly did an experiment and flipped a penny nine times. The first three times the penny turned up tails, the next three times the penny turned up heads, and on the last three flips the penny turned up tails. Kelly saw the pattern and concluded that on the next flip the penny would be most likely to turn up heads.</td>
<td>Strong</td>
</tr>
<tr>
<td>Last year, a large percentage of car crashes were caused by the driver using a cell phone. Also, surveys find that most drivers admit that they are distracted while driving and using their phones. Therefore, using cell phones while driving is dangerous.</td>
<td>Strong</td>
</tr>
<tr>
<td>A friend of yours on Facebook posts that an inventor has created a new technology that when attached to the gas line of a car will double your gas mileage. The inventor says that the reason that the device hasn’t been available before now is because major gasoline companies have prevented the information from getting to the public. You decide that the device probably works.</td>
<td>Strong</td>
</tr>
</tbody>
</table>

**Thank you for your work!**

127
Appendix B. The Lawson Classroom Test of Scientific Reasoning

1. Suppose you are given two clay balls of equal size and shape. The two clay balls also weigh the same. One ball is flattened into a pancake-shaped piece. Which of these statements is correct?
   a. The pancake-shaped piece weighs more than the ball
   b. The two pieces still weigh the same
   c. The ball weighs more than the pancake shaped piece

2. because
   a. The flattened piece covers a larger area.
   b. The ball pushes down more on one spot.
   c. When something is flattened it loses weight.
   d. Clay has not been added or taken away.
   e. When something is flattened it gains weight.

3. To the right are drawings of two cylinders filled to the same level with water. The cylinders are identical in size and shape.

Also shown at the right are two marbles, one glass and one steel. The marbles are the same size but the steel one is much heavier than the glass one.

When the glass marble is put into Cylinder 1 it sinks to the bottom and the water level rises to the 6th mark. If we put the steel marble into Cylinder 2, the water will rise
   a. to the same level as it did in Cylinder 1
   b. to a higher level than it did in Cylinder 1
   c. to a lower level than it did in Cylinder 1
4. because

a. the steel marble will sink faster.
b. the marbles are made of different materials.
c. the steel marble is heavier than the glass marble.
d. the glass marble creates less pressure.
e. the marbles are the same size

5. To the right are drawings of a wide and a narrow cylinder. The cylinders have equally spaced marks on them. Water is poured into the wide cylinder up to the 4th mark (see A). This water rises to the 6th mark when poured into the narrow cylinder (see B).

Both cylinders are emptied (not shown) and water is poured into the wide cylinder up to the 6th mark. How high would this water rise if it were poured into the empty narrow cylinder?

a. To about 8
b. To about 9
c. To about 10
d. To about 12
e. None of these answers is correct

6. because

a. the answer can not be determined with the information given.
b. it went up 2 more before, so it will go up 2 more again.
c. it goes up 3 in the narrow for every 2 in the wide.
d. the second cylinder is narrower.
e. one must actually pour the water and observe to find out.
7. Water is now poured into the narrow cylinder (described in Item 5 above) up to the 11th mark. How high would this water rise if it were poured into the empty wide cylinder?

a. to about 7 ½ 
b. to about 9 
c. to about 8 
d. to about 7 1/3 
e. none of these answers is correct

8. because

a. the ratios must stay the same.
b. one must actually pour the water and observe to find out.
c. the answer cannot be determined with the information given.
d. it was 2 less before so it will be 2 less again.
e. you subtract 2 from the wide for every 3 from the narrow

9. At the right are drawings of three strings hanging from a bar. The three strings have metal weights attached to their ends. String 1 and String 3 are the same length. String 2 is shorter. A 10 unit weight is attached to the end of String 1. A 10 unit weight is also attached to the end of String 2. A 5 unit weight is attached to the end of String 3. The strings (and attached weights) can be swung back and forth and the time it takes to make a swing can be timed.

Suppose you want to find out whether the length of the string has an effect on the time it takes to swing back and forth. Which strings would you use to find out?

a. Only one string 
b. All three strings 
c. 2 and 3 
d. 1 and 3 
e. 1 and 2
10. *because*

a. you must use the longest strings.

b. you must compare strings with both light and heavy weights.

c. only the lengths differ.

d. to make all possible comparisons.

e. the weights differ.

11. Twenty fruit flies are placed in each of four glass tubes. The tubes are sealed. Tubes I and II are partially covered with black paper; Tubes III and IV are not covered. The tubes are placed as shown. Then they are exposed to red light for five minutes. The number of flies in the uncovered part of each tube is shown in the drawing.

![Diagram of fruit flies and light exposure]

This experiment shows that flies respond to (respond means move to or away from):

a. red light but not gravity

b. gravity but not red light

c. both red light and gravity

d. neither red light nor gravity
12. *because*

a. most flies are in the upper end of Tube III but spread about evenly in Tube II.
b. most flies did not go to the bottom of Tubes I and III.
c. the flies need light to see and must fly against gravity.
d. the majority of flies are in the upper ends and in the lighted ends of the tubes.
e. some flies are in both ends of each tube.

13. In a second experiment, a different kind of fly and blue light was used. The results are shown in the drawing.

![Diagram of fly behavior in tubes with blue light and gravity](image_url)

*These data show that these flies respond to* (respond means move to or away from):

a. blue light but not gravity
b. gravity but not blue light
c. both blue light and gravity
d. neither blue light nor gravity

14. *because*

a. some flies are in both ends of each tube.
b. the flies need light to see and must fly against gravity.
c. the flies are spread about evenly in Tube IV and in the upper end of Tube III.
d. most flies are in the lighted end of Tube II but do not go down in Tubes I and III.
e. most flies are in the upper end of Tube I and the lighted end of Tube II.
15. Six square pieces of wood are put into a cloth bag and mixed about. The six pieces are identical in size and shape, however, three pieces are red and three are yellow. Suppose someone reaches into the bag (without looking) and pulls out one piece. What are the chances that the piece is red?

a. 1 chance out of 6 
b. 1 chance out of 3 
c. 1 chance out of 2 
d. 1 chance out of 1 
e. cannot be determined

16. because

a. 3 out of 6 pieces are red. 
b. there is no way to tell which piece will be picked. 
c. only 1 piece of the 6 in the bag is picked. 
d. all 6 pieces are identical in size and shape. 
e. only 1 red piece can be picked out of the 3 red pieces.

17. Three red square pieces of wood, four yellow square pieces, and five blue square pieces are put into a cloth bag. Four red round pieces, two yellow round pieces, and three blue round pieces are also put into the bag. All the pieces are then mixed about. Suppose someone reaches into the bag (without looking and without feeling for a particular shape piece) and pulls out one piece.

What are the chances that the piece is a red round or blue round piece?

a. Cannot be determined 
b. 1 chance out of 3 
c. 1 chance out of 21 
d. 15 chances out of 21 
e. 1 chance out of 2
18. *because*

a. 1 of the 2 shapes is round.
b. 15 of the 21 pieces are red or blue.
c. there is no way to tell which piece will be picked.
d. only 1 of the 21 pieces is picked out of the bag.
e. 1 of every 3 pieces is a red or blue round piece

19. Farmer Brown was observing the mice that live in his field. He discovered that all of them were either fat or thin. Also, all of them had either black tails or white tails. This made him wonder if there might be a link between the size of the mice and the color of their tails. So he captured all of the mice in one part of his field and observed them. Below are the mice that he captured.

![Mice illustration]

*Do you think there is a link between the size of the mice and the color of their tails?*

a. appears to be a link
b. appears not to be a link
c. cannot make a reasonable guess
20. *because*

a. there are some of each kind of mouse.
b. there may be a genetic link between mouse size and tail color.
c. there were not enough mice captured.
d. most of the fat mice have black tails while most of the thin mice have white tails.
e. as the mice grew fatter, their tails became darker

21. The figure below at the left shows a drinking glass and a burning birthday candle stuck in a small piece of clay standing in a pan of water. When the glass is turned upside down, put over the candle, and placed in the water, the candle quickly goes out and water rushes up into the glass (as shown at the right).

![Diagram of glass and candle experiment]

This observation raises an interesting question: Why does the water rush up into the glass?

Here is a possible explanation. The flame converts oxygen into carbon dioxide. Because oxygen does not dissolve rapidly into water but carbon dioxide does, the newly formed carbon dioxide dissolves rapidly into the water, lowering the air pressure inside the glass.

Suppose you have the materials mentioned above plus some matches and some dry ice (dry ice is frozen carbon dioxide). *Using some or all of the materials, how could you test this possible explanation?*

a. Saturate the water with carbon dioxide and redo the experiment noting the amount of water rise.
b. The water rises because oxygen is consumed, so redo the experiment in exactly the same way to show water rise due to oxygen loss.

c. Conduct a controlled experiment varying only the number of candles to see if that makes a difference.

d. Suction is responsible for the water rise, so put a balloon over the top of an open-ended cylinder and place the cylinder over the burning candle.

e. Redo the experiment, but make sure it is controlled by holding all independent variables constant; then measure the amount of water rise.

22. What result of your test (mentioned in #21 above) would show that your explanation is probably wrong?

a. The water rises the same as it did before
b. The water rises less than it did before
c. The balloon expands out
d. The balloon is sucked in

23. A student put a drop of blood on a microscope slide and then looked at the blood under a microscope. As you can see in the diagram below, the magnified red blood cells look like little round balls. After adding a few drops of salt water to the drop of blood, the student noticed that the cells appeared to become smaller.

![Diagram of red blood cells]

This observation raises an interesting question: Why do the red blood cells appear smaller?

Here are two possible explanations: I. Salt ions (Na+ and Cl-) push on the cell membranes and make the cells appear smaller. II. Water molecules are attracted to the salt ions so the water molecules move out of the cells and leave the cells smaller.
To test these explanations, the student used some salt water, a very accurate weighing device, and some water-filled plastic bags, and assumed the plastic behaves just like red-blood-cell membranes. The experiment involved carefully weighing a water-filled bag, placing it in a salt solution for ten minutes and then reweighing the bag.

*What result of the experiment would best show that explanation I is probably wrong?*

a. the bag loses weight  
 b. the bag weighs the same  
 c. the bag appears smaller

24. *What result of the experiment would best show that explanation II is probably wrong?*

a. The bag loses weight  
 b. The bag weighs the same  
 c. The bag appears smaller
Appendix C. Open Response Questionnaire

Please fill in your response to each of the questions in the space provided below each question. Thank you for your time.

1. Tom, Jerry, and Dan are good friends and go fishing together most weekends. They often use the same type of fishing tools and have similar skills in fishing (that is, they often each catch a similar number of fish every time). On their last fishing trip, they had selections of different fishing rods and fishhooks, and they each picked a different location to fish (see the conditions given in the table below). They fished for a total of two hours and the number of fish they each caught during this period is shown below.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Tom</th>
<th>Jerry</th>
<th>Dan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fishing rods</td>
<td>Long</td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td>Fishhooks</td>
<td>Thick</td>
<td>Thin</td>
<td>Thin</td>
</tr>
<tr>
<td>Locations</td>
<td>Point A</td>
<td>Point A</td>
<td>Point B</td>
</tr>
<tr>
<td>Number of fish caught during the two-hour</td>
<td>15</td>
<td>15</td>
<td>8</td>
</tr>
</tbody>
</table>

a) What can you say about the link between each of the three conditions and the number of fish caught?

b) Why do you say that?
The next three questions are about the Canary Island Lizards. Please read the following information (below) about the lizards:

"The Canary Islands are seven islands just west of the African continent. The islands gradually became colonized with life: plants, lizards, birds, etc. Three different species of lizards found on the islands are similar to one species found on the African continent (Thorpe & Brown, 1989). Because of this, scientists assume that the lizards traveled from Africa to the Canary Islands by floating on tree trunks washed out to sea” (Anderson, Fisher, & Norman, 2002, p. 974).

2. Fitness is a term often used by biologists to explain the evolutionary success of certain organisms. Below are description of four fictional female lizards.

<table>
<thead>
<tr>
<th></th>
<th>Lizard A</th>
<th>Lizard B</th>
<th>Lizard C</th>
<th>Lizard D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body length</td>
<td>20 cm</td>
<td>12 cm</td>
<td>10 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>Offspring surviving to adulthood</td>
<td>19</td>
<td>28</td>
<td>22</td>
<td>26</td>
</tr>
<tr>
<td>Age at death</td>
<td>4 years</td>
<td>5 years</td>
<td>4 years</td>
<td>6 years</td>
</tr>
<tr>
<td>Comments</td>
<td>Lizard A is very healthy, strong, and clever</td>
<td>Lizard B has mated with many lizards</td>
<td>Lizard C is dark colored and very quick</td>
<td>Lizard D has the largest territory of all the lizards</td>
</tr>
</tbody>
</table>

a) Which lizard might a biologist consider to be the “most fit”?

b) Why do you say that?
3. a) According to the theory of natural selection, where did the variations in body size in the three species of lizards most likely come?

b) Please explain how and why you came to that conclusion.
4. a) What could cause one species to change into three species over time?

b) Please explain how and why you came to that conclusion.
Appendix D. Interview Protocol

Background Information on Interviewee

What are you studying?

Have you taken science courses in the past (e.g. Biology, Physics, Chemistry)?

How many years have you been at this college?

What do you like most about science?

Question 1

-What connections are there between the fishing rods, fishhooks, and locations?
-What drew your attention to that connection?
-What other factors might you have considered?

Question 2

-Which lizard did you select to be most fit?
-Why do you say that?
-How did you eliminate the other options?
-What data did you pay the most attention to? Why?

If answered Lizard A, ask why Lizards B, C, and D would not be considered the most fit.
*If answered Lizard B, ask why Lizards A, C, and D would not be considered the most fit.
If answered Lizard C, ask why Lizards A, B, and D would not be considered the most fit.
If answered Lizard D, ask why Lizards, A, B, and C would not be considered the most fit.

-How would you persuade someone who chose [any of the alternative responses] that yours is correct?
-What science knowledge have you learned that would support your response?
-What personal experience do you have that would support your response?
-What could someone who chose [any of the alternative responses] say to persuade you that theirs is the correct answer?

Question 3

-From where do you think variations in body size of the three species of lizards came?
Why do you say that?

Jamie’s thought is that: The lizards needed to change in order to survive, so beneficial new traits developed. Is this similar to your response? Why or why not?
  • If not, what is dissimilar about her response? How or why do you know that? How would you convince Jamie of your response?

Lisa’s thought is that: The lizards wanted to become different in size, so beneficial new traits gradually appeared in the population. Is this similar to your response? Why or why not?
  • If not, what is dissimilar about her response? How or why do you know that? How would you convince Lisa of your response?

*John’s thought is that: Random genetic changes and sexual recombination both created new variations. Is this similar to your response? Why or why not?
  • If not, what is dissimilar about his response? How or why do you know that? How would you convince John of your response?

Karl’s thought is that: The island environment caused genetic changes in the lizards. Is this similar to your response? Why or why not?
  • If not, what is dissimilar about his response? How or why do you know that? How could you convince Karl of your response?

Question 4

What would cause one species to change into three species over time?

Jamie’s thought is that: Groups of lizards encountered different island environments so the lizards needed to become new species with different traits in order to survive. Is this similar to your response? Why or why not?
  • If not, what is dissimilar about her response? How or why do you know that? How would you convince Jamie of your response?

Lisa’s thought is that: Groups of lizards must have been geographically isolated from other groups and random genetic changes must have accumulated in these lizard populations over time. Is this similar to your response? Why or why not?
  • If not, what is dissimilar about her response? How or why do you know that? How would you convince Lisa of your response?

John’s thought is that: There may be minor variations, but all lizards are essentially alike and all are members of a single species. Is this similar to your response? Why or why not?
  • If not, what is dissimilar about his response? How or why do you know that? How would you convince John of your response?
Karl’s thought is that: In order to survive, different groups of lizards needed to adapt to the different islands, and so all organisms in each group gradually evolved to become a new lizard species. Is this similar to your response? Why or why not?

- If not, what is dissimilar about his response? How or why do you know that? How could you convince Karl of your response?

*italicized responses are correct